

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 4, 2002, 10:28:22 ; Search time 43.55 Seconds  
(without alignments)  
55.613 Million cell updates/sec

Title: US-09-822-698A-28  
Perfect score: 54  
Sequence: 1 XXHTCGVWXPXXX 14

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_19:\*\*

- 1: sp\_archaea:\*\*
- 2: sp\_bacteria:\*\*
- 3: sp\_fungi:\*\*
- 4: sp\_human:\*\*
- 5: sp\_invertebrate:\*\*
- 6: sp\_mammal:\*\*
- 7: sp\_mhc:\*\*
- 8: sp\_organelle:\*\*
- 9: sp\_phase:\*\*
- 10: sp\_plant:\*\*
- 11: sp\_rodent:\*\*
- 12: sp\_virus:\*\*
- 13: sp\_vertebrate:\*\*
- 14: sp\_unclassified:\*\*
- 15: sp\_virus:\*\*
- 16: sp\_bacteriap:\*\*
- 17: sp\_archaeap:\*\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	74.1	1077	081441	081441 arabidopsis
2	37	68.5	124	098EW3	Q98EW3 rhizobium l
3	37	68.5	301	09CCT9	Q9CCT9 mycobacteri
4	37	68.5	306	007152	007152 mycobacteri
5	37	68.5	461	091197	P91197 caenorhabdi
6	37	68.5	716	09FC94	Q9FC94 xylella fas
7	36	66.7	158	09F8J1	Q9F8J1 carboxydoth
8	36	66.7	421	098LM4	Q98LM4 rhizobium l
9	36	66.7	422	09L011	Q9L011 streptomyce
10	36	66.7	540	046134	046134 locusta mig
11	36	66.7	552	09VC74	Q9VC74 drosophila
12	36	66.7	581	096WM9	Q96WM9 botrytis ci
13	36	66.7	762	09RQ15	Q9RQ15 neisseria d
14	35	64.8	329	09RYX0	Q9RYX0 deinococcus
15	35	64.8	448	09KGB9	Q9KGB9 bacillus ha
16	35	64.8	710	014407	014407 neurospora

17	34	63.0	168	5	Q9V5J1	Q9V5J1 drosophila
18	34	63.0	263	16	Q92NL0	Q92NL0 rhizobium m
19	34	63.0	285	16	Q9C168	Q9C168 lactococcus
20	34	63.0	296	5	Q9GSF1	Q9GSF1 branchiost
21	34	63.0	308	10	Q94KT0	Q94KT0 zea mays (m
22	34	63.0	401	15	011575	011575 human immun
23	34	63.0	408	15	011574	011574 human immun
24	34	63.0	435	5	Q19374	Q19374 caenorhabdi
25	34	63.0	441	17	Q9HQN2	Q9HQN2 halobacteri
26	34	63.0	545	5	Q96G31	Q96G31 heliothis v
27	34	63.0	726	2	Q9ACZ7	Q9ACZ7 streptomyce
28	34	63.0	792	4	Q9H0D2	Q9H0D2 homo sapien
29	34	63.0	812	16	Q911X1	Q911X1 pseudomonas
30	34	63.0	824	5	Q966U9	Q966U9 drosophila
31	34	63.0	825	2	P82593	P82593 streptomyce
32	34	63.0	1596	13	Q918E1	Q918E1 fugu rubrip
33	33	61.1	199	12	Q91GW5	Q91GW5 foot-and-mo
34	33	61.1	206	12	Q83476	Q83476 maize strea
35	33	61.1	206	12	Q56970	Q56970 maize strea
36	33	61.1	206	12	Q73468	Q73468 maize strea
37	33	61.1	206	12	Q73472	Q73472 maize strea
38	33	61.1	206	12	Q73474	Q73474 maize strea
39	33	61.1	210	16	Q9KEF3	Q9KEF3 bacillus ha
40	33	61.1	211	2	Q04904	Q04904 escherichia
41	33	61.1	252	4	Q9HAI3	Q9HAI3 homo sapien
42	33	61.1	252	4	Q96A62	Q96A62 homo sapien
43	33	61.1	260	16	Q9RVR6	Q9RVR6 deinococcus
44	33	61.1	294	5	Q9NFP4	Q9NFP4 cuplennius
45	33	61.1	302	10	Q82593	Q82593 arabidopsis

#### ALIGNMENTS

RESULT 1  
081441  
ID 081441 PRELIMINARY; PRT; 1077 AA.  
AC 081441;  
DT 01-NOV-1998 (TREMblrel. 08, Created)  
DT 01-NOV-1998 (TREMblrel. 08, Last sequence update)  
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)  
DE T24H24.17 PROTEIN (PUTATIVE REVERSE TRANSCRIPTASE).  
GN T24H24.17 OR AT4G04000.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopses.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. COLUMBIA;  
RA WASHU;  
RT "The A. thaliana Genome Sequencing Project.";  
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. COLUMBIA;  
RA Courtney L., Stoneking T., Langston Y., Mead K.;  
RT "The sequence of A. thaliana T24H24.";  
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. COLUMBIA;  
RA Waterston R.;  
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RA Lamar B., Stoneking T., Stumpf J., Mewes H.W., Lampke K.,  
RT Mayer K.F.X.;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE FROM N.A.  
RA EU Arabidopsis sequencing project;

```

RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF075598; AAC28197.1; -.
DR EMBL: AL161499; CAB77868.1; -.
DR InterPro: IPR000477; RVTse.
DR InterPro: IPR000531; TonB_boxC.
DR Pfam: PF000078; rvt_2.
DR PROSITE: PS00430; TONB_DEPENDENT_REC_1; UNKNOWN_1.
KW RNA-directed DNA polymerase.
SQ SEQUENCE 1077 AA; 121858 MW; B272175E4BF57D3A CRC64;

Query Match 74.1%; Score 40; DB 10; Length 1077;
Best Local Similarity 55.6%; Pred. No. 35;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 HTGXGVWXP 11
   |||||
Db 959 HAGSGIWP 967

RESULT 2
ID Q98EW3 PRELIMINARY; PRT; 124 AA.
AC Q98EW3;
DT 01-OCT-2001 (TReMBLrel. 18, Created)
DT 01-OCT-2001 (TReMBLrel. 18, Last sequence update)
DE DEGRADATION PROTEIN.
GN MLL4053.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti.";
RL DNA Res. 7:331-338(2000).
DR EMBL: AP003003; BAB50804.1; -.
KW Complete proteome.
SQ SEQUENCE 124 AA; 13315 MW; CF535F335A73D135 CRC64;

Query Match 68.5%; Score 37; DB 16; Length 124;
Best Local Similarity 55.6%; Pred. No. 13;
Matches 5; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 HTGXGVWXP 11
   |||||
Db 14 HAGEGAWTP 22

RESULT 3
ID Q9CCT9 PRELIMINARY; PRT; 301 AA.
AC Q9CCT9;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DE PUTATIVE OXIDOREDUCTASE.
GN MLO458.
OS Mycobacterium leprae.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1769;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TN;
RX MEDLINE=21128732; PubMed=11234002;
RA Cole S.T., Eiglmeier K., Parkhill J., James K.D., Thomson N.R.,
RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,
RA Holroyd S., Hornsby T., Jagers K., Lacroix C., Maclean J., Moule S.,
RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
RA Rutter S., Seeger K., Simon S., Simmonds M., Skellton J., Squares R.,
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA Barrell B.G.;
RT "Massive gene decay in the leprosy bacillus.";
RL Nature 409:1007-1011(2001).
DR EMBL: AL583918; CAC29966.1; -.
DR Leproma; MLO458; -.
DR InterPro: IPR001395; Aldo_ket_red.
DR Pfam: PF00248; aldo_ket_red.2.
DR PRINTS: PR00069; ALDKETREDTASE.
KW Complete proteome.
SQ SEQUENCE 301 AA; 32874 MW; 53D5D2F57744AFF9 CRC64;

Query Match 68.5%; Score 37; DB 16; Length 301;
Best Local Similarity 75.0%; Pred. No. 32;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 TGXGVWXP 11
   |||||
Db 45 TGKGVWGP 52

RESULT 4
ID O07152 PRELIMINARY; PRT; 306 AA.
AC O07152;
DT 01-JUL-1997 (TReMBLrel. 04, Created)
DT 01-JUL-1997 (TReMBLrel. 04, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE PUTATIVE OXIDOREDUCTASE.
GN MCL581.18C.
OS Mycobacterium leprae.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1769;
RN [1]
RP SEQUENCE FROM N.A.
RC Oliver K., Harris D.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC MEDLINE=93188700; PubMed=8446027;
RX Eiglmeier K., Honore N., Woods S.A., Caudron B., Cole S.T.;
RT "Use of an ordered cosmid library to deduce the genomic organization
RT of Mycobacterium leprae.";
RL Mol. Microbiol. 7:197-206(1993).
DR EMBL: Z96801; CAB09639.1; -.
DR InterPro: IPR001395; Aldo_ket_red.
DR Pfam: PF00248; aldo_ket_red.2.
SQ SEQUENCE 306 AA; 33463 MW; CA3781FE6338B17A CRC64;

Query Match 68.5%; Score 37; DB 2; Length 306;
Best Local Similarity 75.0%; Pred. No. 33;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 TGXGVWXP 11
   |||||
Db 50 TGKGVWGP 57

```

```

RESULT 5
P91197 PRELIMINARY; PRT; 461 AA.
ID P91197
AC P91197
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE HYPOHETICAL 52.7 KDA PROTEIN.
GN D2092.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OC NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RL investigating biology. The C. elegans Sequencing Consortium.";
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Gattung S., Magg L.;
RT "The sequence of C. elegans cosmid D2092.";
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.
DR EMBL; U88167; ABA4223.1; -.
DR InterPro; IPR001188; GABAA_receptor.
DR InterPro; IPR001175; Neur_chan_I.
DR Pfam; PF02931; Neur_chan_I.
DR Pfam; PF02932; Neur_chan_I.
DR PRINTS; PS00252; NRIONCHANNEL.
DR PROSITE; PS00236; NEUROTR_ION_CHANNEL; 1.
DR Glycoprotein; Hypothetical protein; Ionic channel;
KW Postsynaptic membrane; Transmembrane.
SQ SEQUENCE 461 AA; 52718 MW; 6182A7F827357B92 CRC64;

Query Match 68.5%; Score 37; DB 5; Length 461;
Best Local Similarity 66.7%; Pred. No. 51;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 HTGXGVWXP 11
Db 136 HTGTVVWIP 144

RESULT 6
Q9PC94 PRELIMINARY; PRT; 716 AA.
ID Q9PC94
AC Q9PC94
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE HYPOHETICAL PROTEIN XF1887.
GN XF1887.
OS Xylella fastidiosa.
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
OC Xylella.
OC NCBI_TaxID=2371;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN=9A5C;
RX MEDLINE=20365717; PubMed=10910347;
RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA Alvarenga R., Alves L.M.C., Araya J.E., Bala G.S., Baptista C.S.,
RA Barro M.R.P., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA Buato M.R.P., Camargo A.A., Camargo L.E.A., Carreira D.M., Carreira H.,
RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docere J., El-Dorry H.,
RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro L.R.,
RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furian L.R.,
RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S., Gruber A.,
RA Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E., Kitajima J.P.,
RA Krieger J.E., Kuramae E.E., Laigret F., Lambais H.R., Leite L.C.C.,
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C., Machado J.A.,
RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Macino C.L.,
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto C.E.S.,
RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Tesquero J.B.,
RA Queglio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA da Silveira J.F., Silvestri M.L.Z., Silveira W.J., de Souza A.A.,
RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tshako M.H.,
RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
RT "The genome sequence of the plant pathogen Xylella fastidiosa.";
RL Nature 406:151-159(2000).
DR EMBL; AE004008; AAF84693.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 716 AA; 79375 MW; EAF086E2315BBD5C CRC64;

Query Match 68.5%; Score 37; DB 16; Length 716;
Best Local Similarity 55.6%; Pred. No. 82;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 3 HTGXGVWXP 11
Db 21 HAGEGMWVP 29

RESULT 7
Q9F8J1 PRELIMINARY; PRT; 158 AA.
ID Q9F8J1
AC Q9F8J1
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE PHOSPHATIDYL SERINE DECARBOXYLASE (FRAGMENT).
OS Carboxydotherrnus hydrogenoformans.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Carboxydotherrnus.
OX NCBI_TaxID=129958;
RN [1]
RP SEQUENCE FROM N.A.
RA Gonzalez J.M., Robb F.T.;
RT "A genomic survey of the extreme thermophilic, CO-utilizing bacterium
RL Carboxydotherrnus hydrogenoformans.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF244642; AAG23591.1; -.
DR InterPro; IPR003817; PS_Dcarbxylyase.
DR Pfam; PF02666; PS_Dcarbxylyase; 1.
FT NON_TER 1 158
FT NON_TER 158 158
SQ SEQUENCE 158 AA; 17687 MW; 9C2669923FF7898C CRC64;

Query Match 66.7%; Score 36; DB 2; Length 158;
Best Local Similarity 75.0%; Pred. No. 25;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 4 TGXGVWXP 11
Db 136 TGHGVWMP 143

RESULT 8
Q98LM4 PRELIMINARY; PRT; 421 AA.
AC Q98LM4;
DT 01-OCT-2001 (TREMBLrel. 18, Created)
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
DE N-CARBAMYL-L-AMINO ACID AMIDOHYDROLASE.
GN MLR0967.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpō S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti.";
RL DNA Res. 7:331-338(2000).
DR EMBL: AP002996; BAB48439.1; -.
DR InterPro: IPR000515; BPD.transp.
DR PROSITE; PS00402; BPD_TRANSF_INN_MEMBER; UNKNOWN_1.
DR Hydrolase; Complete proteome.
KW
SQ SEQUENCE 421 AA; 45172 MW; 67640F33C747287E CRC64;

Query Match 66.7%; Score 36; DB 16; Length 421;
Best Local Similarity 71.4%; Pred. No. 72;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 HTXGVW 9
Db 33 HSGTGW 39

RESULT 9
Q9L011 PRELIMINARY; PRT; 422 AA.
AC Q9L011;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE HYPOTHETICAL 43.2 KDA PROTEIN.
GN SCC30.10.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Seeger K.J., Harris D.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Cerdano A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RX MEDLINE=97000351; PubMed=8843436;

```

```

RA Redenbach M., Kieser H.M., Denapate D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL MOL. Microbiol. 21:77-96(1996).
DR EMBL; AL352972; CAB88180.1; -.
KW Hypothetical protein.
SQ SEQUENCE 422 AA; 43245 MW; 4020E69070FB6417 CRC64;

```

```

Query Match 66.7%; Score 36; DB 2; Length 422;
Best Local Similarity 62.5%; Pred. No. 72;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 4 TGXGVWXP 11
Db 313 TGAGLWTP 320

```

```

RESULT 10
O46134 PRELIMINARY; PRT; 540 AA.
AC O46134;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE NICOTINIC ACETYLCHOLINE RECEPTOR, ALPHA3 SUBUNIT (FRAGMENT).
GN NACHR.
OS Locusta migratoria (Migratory locust).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Orthopteroidea; Orthoptera; Caelifera;
OC Acridoidea; Acridoidea; Acrididae; Locusta.
OX NCBI_TaxID=7004;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98325051; PubMed=9660807;
RA Hermen B., Stetzer E., Thees R., Heiermann R., Schrattenholz A.,
RA Ebdinghaus U., Kretschmer A., Methfessel C., Reinhardt S.,
RA Maelicke A.;
RT "Neuronal nicotinic receptors in the locust Locusta migratoria.";
RL J. Biol. Chem. 273:18394-18404(1998).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.
DR EMBL; AJ000392; CAA04054.1; -.
DR InterPro: IPR00188; GABAA_receptor.
DR InterPro: IPR001175; Neur_chan.
DR Pfam; PF02931; Neur_chan_LBD; 1.
DR Pfam; PF02932; Neur_chan_memb; 1.
DR PRINTS; PR00252; NRIONCHANNEL.
DR PROSITE; PS00236; NEUROTR_ION_CHANNEL; 1.
KW Glycoprotein; Ionic channel; Postsynaptic membrane; Receptor;
KW Transmembrane.
FT NON_TER 1
SQ SEQUENCE 540 AA; 60201 MW; 809EB7E98369DFFA CRC64;

```

```

Query Match 66.7%; Score 36; DB 5; Length 540;
Best Local Similarity 66.7%; Pred. No. 94;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 3 HTXGVWXP 11
Db 101 HTXGVWKP 109

```

```

RESULT 11
Q9VC74 PRELIMINARY; PRT; 552 AA.
AC Q9VC74;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE NACRALPHA-96AA PROTEIN.

```



GN NACR-ALPHA-96AA OR NACRALPHA-96AA OR CG5610.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BERKELEY;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Peiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Bencos P.V., Bernan B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,  
 RA Burdick K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleeb J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster."  
 RL Science 287:2185-2195(2000).  
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).  
 CC -!- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.  
 DR EMBL: AF003747; AAF56301.1;  
 DR FlyBase: FBgn0000036; nacr-alpha-96Aa.  
 DR InterPro: IPR000188; GABA<sub>A</sub> receptor.  
 DR InterPro: IPR001175; Neur\_chan.  
 DR Pfam: PF02931; Neur\_chan\_LBD; 2.  
 DR Pfam: PF02932; Neur\_chan\_memb; 1.  
 DR PRINTS: PR00252; NRIONCHANNEL.  
 DR PROSITE: PS00236; NEUROTRP\_ION\_CHANNEL; 1.  
 KW Glycoprotein; Ionic channel; Postsynaptic membrane; Transmembrane.  
 SQ SEQUENCE 552 AA; 62218 MW; DABAIC31654DF636 CRC64;

Query Match 66.7%; Score 36; DB 5; Length 552;  
 Best Local Similarity 66.7%; Pred. No. 96;  
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 HTGXGVWXP 11  
 ||| |||  
 DB 118 HTGKVVWKP 126

RESULT 12  
 Q96WM9

ID Q96WM9 PRELIMINARY; PRT; 581 AA.  
 AC Q96WM9;  
 DT 01-DEC-2001 (T-EMBLrel. 19, Created)  
 DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)  
 DT 01-DEC-2001 (T-EMBLrel. 19, Last annotation update)  
 DE LACCASE 2 (EC 1.10.3.2).  
 GN LCC2.  
 OS Botrytis cinerea (Botryotinia fuckeliana).  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Leotiomycetes;  
 OC Helotiales; Sclerotiniaceae; Botryotinia.  
 OX NCBI\_TaxID=40559;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=SAS56;  
 RA Schouten A., Vermeer J.E.M., van Kan J.A.L.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF243855; AAK77953.1;  
 KW Oxidoreductase.  
 SQ SEQUENCE 581 AA; 63434 MW; 674947DAFD6BC757 CRC64;  
 Query Match 66.7%; Score 36; DB 3; Length 581;  
 Best Local Similarity 62.5%; Pred. No. 1e+02;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 4 TGXGVWXP 11  
 ||| |||  
 DB 458 TGFGIWHP 465  
 RESULT 13  
 Q9RQI5 PRELIMINARY; PRT; 762 AA.  
 ID Q9RQI5;  
 DT 01-MAY-2000 (T-EMBLrel. 13, Created)  
 DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)  
 DT 01-DEC-2001 (T-EMBLrel. 19, Last annotation update)  
 DE 1,4-ALPHA-GLUCAN BRANCHING ENZYME (EC 2.4.1.18).  
 GN GLGB.  
 OS Neisseria denitrificans.  
 OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.  
 OX NCBI\_TaxID=494;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC14686;  
 RX MEDLINE=95337488; PubMed=10407163;  
 RA Buettcher V., Quanz M., Willmitzer L.;  
 RT "Molecular cloning, functional expression and purification of a glucan  
 RT branching enzyme from Neisseria denitrificans."  
 RL Biochim. Biophys. Acta 1432:406-412(1999).  
 DR EMBL: AF102867; AAF04747.1;  
 DR InterPro: IPR000461; Alpha\_amyase.  
 DR InterPro: IPR004193; isomylase\_N.  
 DR Pfam: PF00128; alpha-amyase; 1.  
 DR Pfam: PF02922; isomylase\_N; 1.  
 KW Transferrase; Glycosyltransferase.  
 SQ SEQUENCE 762 AA; 86367 MW; 9922F80406DDE3E3 CRC64;

Query Match 66.7%; Score 36; DB 2; Length 762;  
 Best Local Similarity 71.4%; Pred. No. 1.4e+02;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 HTGXGVW 9  
 ||| |||  
 DB 181 HTGNGIWH 187

RESULT 14  
 Q9RYX0 PRELIMINARY; PRT; 329 AA.  
 ID Q9RYX0;  
 AC Q9RYX0;  
 DT 01-MAY-2000 (T-EMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)  
DE PYRIDOXAMINE KINASE.  
GN DRA0184.  
OS Deinococcus radiodurans.  
OC Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.  
OX NCBI\_TaxID=1299;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=RI;  
RX MEDLINE=20336896; PubMed=10567266;  
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,  
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,  
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,  
RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zalewski C.,  
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,  
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,  
RA Fraser C.M.;  
RT "Genome sequence of the radioresistant bacterium Deinococcus  
RT radiodurans R1.";  
RL Science 286.1571-1577(1999).  
DR EMBL; AE001862; AAF12189.1; -  
DR TIGR; DRA0184; -  
DR InterPro; IPR002173; PfkB.  
DR pfam; PF00294; pfkB; 1.  
KW Kinase; Complete proteome.  
SQ SEQUENCE 329 AA; 34962 MW; 3C39AC6C42C4CB24 CRC64;

Query Match 64.8%; Score 35; DB 16; Length 329;  
Best Local Similarity 71.4%; Pred. No. 85;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 HTGXGVW 9  
||| |  
DB 82 HTGYGAW 88

RESULT 15  
Q9KGB9  
ID O9KGB9 PRELIMINARY; PRT; 448 AA.  
AC O9KGB9;  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE BH0185 PROTEIN.  
GN BH0185.  
OS Bacillus halodurans.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
OC Bacillus/Staphylococcus group; Bacillus.  
OX NCBI\_TaxID=86665;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C-125 / JCM 9153;  
RX MEDLINE=20512582; PubMed=11058132;  
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,  
RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,  
RA Horikoshi K.;  
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus  
RT halodurans and genomic sequence comparison with Bacillus subtilis.";  
RL Nucleic Acids Res. 28.4317-4331(2000).  
DR EMBL; AF001507; BAB03904.1; -  
DR InterPro; IPR002604; ATZ\_TRZ.  
DR pfam; PF01685; ATZ\_TRZ; 1.  
KW Complete proteome.  
SQ SEQUENCE 448 AA; 50007 MW; 3FF74BE7DE60E6F3 CRC64;

Query Match 64.8%; Score 35; DB 16; Length 448;  
Best Local Similarity 55.6%; Pred. No. 1.2e+02;  
Matches 5; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 HTGXGVWXP 11

Db 220 HIGFTWGP 228  
| | | | |

Search completed: October 4, 2002, 10:28:24  
Job time: 391 sec

---



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 4, 2002, 10:26:16 ; Search time 24.92 Seconds  
(without alignments)  
53.983 Million cell updates/sec

Title: US-09-822-698a-28

Perfect score: 54

Sequence: 1 XXHTGXGVWXPXXX 14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR-71:\*  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	40	74.1	1077	2 T01474	hypothetical prote
2	38	70.4	250	2 A54308	PE-1 protein - hum
3	38	70.4	656	2 I49299	receptor interacti
4	38	70.4	807	2 A32582	protein kinase YAK
5	37	68.5	301	2 B86966	probable oxidoredu
6	37	68.5	461	2 T25671	hypothetical prote
7	37	68.5	557	2 S12359	nicotinic acetylch
8	37	68.5	716	2 G82627	hypothetical prote
9	36	66.7	567	1 AC8FAL	nicotinic acetylch
10	35	64.8	329	2 B75615	pyridoxamine kinas
11	35	64.8	448	2 AB3673	hypothetical prote
12	35	64.8	710	2 T46589	ropy-2 protein [im
13	34	63.0	285	2 B86687	conserved hypothet
14	34	63.0	397	2 S27600	N-acetylglutamate
15	34	63.0	414	2 B98354	hypothetical prote
16	34	63.0	435	2 T20819	hypothetical prote
17	34	63.0	441	2 E84264	isochorismate synt
18	34	63.0	473	2 I54210	N-acetylglactosam
19	34	63.0	550	1 FGRTA	fibrinogen alpha c
20	34	63.0	812	2 AB3379	glycogen phosphory
21	34	63.0	825	2 A59296	alpha-L-arabinofur
22	33	61.1	60	2 JN0251	cytochrome c552 -
23	33	61.1	114	1 G69614	diacylglycerol kin
24	33	61.1	206	2 T10113	hypothetical prote
25	33	61.1	208	2 AC3091	hypothetical prote
26	33	61.1	210	2 C83762	endo-1,4-beta-xyla
27	33	61.1	224	2 G98195	hypothetical prote
28	33	61.1	260	2 C75454	hypothetical prote
29	33	61.1	281	2 AB2288	hypothetical prote

#### ALIGNMENTS

##### RESULT 1

T01474

hypothetical protein T24H24.17 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 24-Mar-1999

C:Accession: T01474

R:Courtney, L.; Stoneking, T.; Langston, Y.; Mead, K.

submitted to the EMBL Data Library, August 1998

A:Description: The sequence of A. thaliana T24H24.

A:Reference number: Z14333

A:Accession: T01474

A:Status: translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-1077 <CD>

A:Cross-references: EMBL:AF075598; NID:g3293581; PID:g3377824

A:Experimental source: cultivar Columbia

C:Genetics:

A:Map position: 4

A:Introns: 172/2; 583/2; 833/1; 978/1

A:Note: T24H24.17

Query Match 74.1%; Score 40; DB 2; Length 1077;  
Best Local Similarity 55.6%; Pred. No. 17;  
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 HTGXGVWXP 11  
DB 959 HAGSGIWL 967

##### RESULT 2

A54308

PE-1 protein - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 24-May-1996 #sequence\_revision 24-May-1996 #text\_change 05-Nov-1999

C:Accession: A54308

R:Klemsz, M.; Hromas, R.; Raskind, W.; Bruno, E.; Hoffman, R.

Genomics 20, 291-294, 1994

A:Title: PE-1, a novel ETS oncogene family member, localizes to chromosome 1q21-q23.

A:Reference number: A54308; MOID:94292214

A:Accession: A54308

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: mRNA

A:Residues: 1-250 <RES>

A:Cross-references: GB:L16464; NID:g291991; PIDN:AAA60949.1; PID:g291992

C:Genetics:

A:Gene: GDB:ETV3; PE-1

A:Cross-references: GDB:362716; OMIM:164873

A:Map position: 1q21-q23

C:Superfamily: ets DNA-binding domain homology

F:59-138/Domain: ets DNA-binding domain homology <ETS>

hypothetical prote  
acyltransferase ho  
conserved hypothet  
NADPH-dependent gl  
hypothetical prote  
transcription fact  
gas-vesicle operon  
probable arylsulfa  
hypothetical prote  
hypothetical prote  
beta-galactosidase  
DNA topoisomerase  
hypothetical prote  
genome polyprotein  
myosin heavy chain

```
Yeast 12, 787-797, 1996
A;Title: Sequence analysis of a 40.7 kb segment from the left arm of yeast chromosome
of chromosome XI.
A;Reference number: S71643; MUID:96408771
A;Accession: S71656
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-807 <KAF>
A;Cross-references: EMBL:X87371; NID:g854542; PIDN:CAA60814.1; PID:g854556
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, May 1995
C;Genetics:
A;Gene: SGD:YAK1
A;Cross-references: SGD:S0003677; MIPS:YJL141C
A;Map position: 10L
C;Function:
C;Description: phosphotransferase; protein kinase; may be involved in cell-cycle regu
C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase hom
C;Keywords: ATP; phosphoprotein; phosphotransferase; serine/threonine-specific protei
F;367-630/Domain: protein kinase homology <KIN>
F;375-383/Region: protein kinase ATP-binding motif

Query Match 70.4%; Score 38; DB 2; Length 807;
Best Local Similarity 55.6%; Pred. No. 30;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 HTXGVWXP 11
|||
DB 43 HMGRIWNP 51

RESULT 5
B86966
probable oxidoreductase M10458 [imported] - Mycobacterium leprae
C;Species: Mycobacterium leprae
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 10-May-2001
C;Accession: B86966
R;Cole, S.T.; Eiglmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.;
R.; Davies, R.M.; Devlin, K.; Duthoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holro
eam, M.A.; Rutherford, K.M.;
Nature 409, 1007-1011, 2001
A;Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.;
A;Title: Massive gene decay in the leprosy bacillus.
A;Reference number: A86909; MUID:21128732; PMID:11234702
A;Accession: B86966
A;Status: Preliminary
A;Molecule type: DNA
A;Residues: 1-301 <STO>
A;Cross-references: GB:AL450380; NID:g13092696; PIDN:CA229966.1; GSPDB:GN00147
C;Genetics:
A;Gene: M10458
C;Superfamily: fission yeast pyridoxine 4-dehydrogenase;

Query Match 68.5%; Score 37; DB 2; Length 301;
Best Local Similarity 75.0%; Pred. No. 17;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 TGXGVWXP 11
|||
DB 45 TGXGVWGP 52

RESULT 6
T25671
hypothetical protein D2092.3 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
C;Accession: T25671
R;Gattung, S.; Maggi, L.
submitted to the EMBL Data Library, February 1997
A;Description: The sequence of C. elegans cosmid D2092.
A;Reference number: Z20067

Query Match 70.4%; Score 38; DB 2; Length 250;
Best Local Similarity 75.0%; Pred. No. 9.1;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 TGXGVWXP 11
|||
DB 223 TGAGWIP 230

RESULT 3
I49299
receptor interacting protein RIP - mouse
C;Species: Mus musculus (house mouse)
C;Date: 09-Mar-1996 #sequence_revision 09-Mar-1996 #text_change 05-Nov-1999
C;Accession: I49299
R;Stanger, B.Z.; Leder, P.; Lee, T.H.; Kim, E.; Seed, B.
Cell 81, 513-523, 1995
A;Title: RIP: a novel protein containing a death domain that interacts with Fas/APO-1 (C
A;Reference number: A56913; MUID:95277838
A;Accession: I49299
A;Status: Preliminary
A;Molecule type: mRNA
A;Residues: 1-656 <RES>
A;Cross-references: EMBL:U25995; NID:g829618; PIDN:AAB60487.1; PID:g829619
C;Genetics:
A;Gene: RIP
C;Superfamily: protein kinase homology
F;15-293/Domain: protein kinase homology <KIN>

Query Match 70.4%; Score 38; DB 2; Length 656;
Best Local Similarity 75.0%; Pred. No. 24;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 TGXGVWXP 11
|||
DB 473 TGXGVWIP 480

RESULT 4
A32582
protein kinase YAK1 (EC 2.7.1.1) - Yeast (Saccharomyces cerevisiae)
N;Alternate names: protein Y0652; protein YJL141C
C;Species: Saccharomyces cerevisiae
C;Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 24-Sep-1999
C;Accession: A32582; S55172; S56923; S71656
R;Garrett, S.; Broach, J.
Genes Dev. 3, 1336-1348, 1989
A;Title: Loss of Ras activity in Saccharomyces cerevisiae is suppressed by disruptions o
A;Reference number: A32582; MUID:90108683
A;Accession: A32582
A;Molecule type: DNA
A;Residues: 1-807 <GAR>
A;Cross-references: GB:X16036; NID:g4795; PIDN:CAA34192.1; PID:g4796
R;Katsoulou, C.; Tzermia, M.; Alexandraki, D.
submitted to the EMBL Data Library, May 1995
A;Description: The complete sequence of a 40.7 kb segment located on the left arm of ye
A;Reference number: S55159
A;Reference number: S55172
A;Accession: S55172
A;Molecule type: DNA
A;Residues: 1-807 <KAT>
A;Cross-references: EMBL:X87371; NID:g854542; PIDN:CAA60814.1; PID:g854556
R;Katsoulou, C.; Tzermia, M.; Alexandraki, D.
submitted to the Protein Sequence Database, September 1995
A;Reference number: S56912
A;Accession: S56923
A;Molecule type: DNA
A;Residues: 1-807 <KAW>
A;Cross-references: EMBL:Z49417; NID:g1015549; PIDN:CAA89437.1; PID:g1015551; MIPS:YJL14
R;Katsoulou, C.; Tzermia, M.; Tavernarakis, N.; Alexandraki, D.
```

A:Accession: T25671  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-461 <GAT>  
 A:Cross-references: EMBL:U88167; PIDN:AA842223.1; GSPDB:GN00019; CESP:D2092.3  
 A:Experimental source: strain Bristol N2; clone D2092  
 C:Genetics:  
 A:Gene: CESP:D2092.3  
 A:Map position: 1  
 A:Introns: 36/1; 80/3; 119/2; 191/1; 243/1; 274/1; 303/1; 382/2  
 C:Superfamily: acetylcholine receptor

Query Match 68.5%; Score 37; DB 2; Length 461;  
 Best Local Similarity 66.7%; Pred. No. 26;  
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 HTGXGVWXP 11  
 ||| |||  
 Db 136 HTGVVWIP 144

RESULT 7  
 S12359  
 nicotinic acetylcholine receptor alpha-L1 chain precursor - desert locust  
 C:Species: Schistocerca gregaria (desert locust)  
 C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 20-Aug-1999  
 C:Accession: S12359  
 R:Marshall, J.; Buckingham, S.D.; Shingai, R.; Lunt, G.G.; Goosey, M.W.; Darlison, M.G.; EMBO J. 9, 4391-4398, 1990  
 A:Title: Sequence and functional expression of a single alpha subunit of an insect nicot  
 A:Reference number: S12359; MUID:91092263  
 A:Accession: S12359  
 A:Molecule type: mRNA  
 A:Residues: 1-557 <MAR>  
 A:Cross-references: EMBL:X55439; NID:g10133; PIDN:CAA39081.1; PID:g10134  
 C:Superfamily: acetylcholine receptor  
 C:Keywords: glycoprotein; ion channel; neurotransmitter receptor; transmembrane protein  
 F:1-23/Domain: signal sequence #status predicted <SIG>  
 F:24-557/Product: nicotinic acetylcholine receptor alpha-L1 chain #status predicted <MAT>  
 F:245-266/Domain: transmembrane #status predicted <TM1>  
 F:274-295/Domain: transmembrane #status predicted <TM2>  
 F:308-329/Domain: transmembrane #status predicted <TM3>  
 F:501-523/Domain: transmembrane #status predicted <TM4>  
 F:47,235/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 68.5%; Score 37; DB 2; Length 557;  
 Best Local Similarity 66.7%; Pred. No. 32;  
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 HTGXGVWXP 11  
 ||| |||  
 Db 135 HTGVVWIP 143

RESULT 8  
 G82627  
 hypothetical protein XF1887 [imported] - Xylella fastidiosa (strain 9a5c)  
 C:Species: Xylella fastidiosa  
 C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000  
 C:Accession: G82627  
 R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen  
 Nature 406, 151-157, 2000  
 A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.  
 A:Reference number: A82515; MUID:20365717  
 A:Note: for a complete list of authors see reference number A59328 below  
 A:Accession: G82627  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-716 <SIM>  
 A:Cross-references: GB:AE004008; GB:AE003849; NID:g9106961; PIDN:AAF84693.1; GSPDB:GN001  
 A:Experimental source: strain 9a5c

R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, V.A.; Acencio, M.; Alvarenga, R.  
 Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrier  
 as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.  
 submitted to Genbank, June 2000  
 A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Fr  
 J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; La  
 chado, M.A.; Madeira, A.M.B.N.; Madeira, H.Y.F.; Marino, C.L.; Marques, M.V.; Martins  
 A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C  
 F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri,  
 Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; do Sa, R.G.; Santelli, R.V.; Sawa  
 A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silv  
 M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjowski-Almeida, S.; Vettore, A.L  
 A:Reference number: A59328  
 A:Contents: annotation  
 C:Genetics:  
 A:Gene: XF1887

Query Match 68.5%; Score 37; DB 2; Length 716;  
 Best Local Similarity 55.6%; Pred. No. 41;  
 Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 HTGXGVWXP 11  
 ||| |||  
 Db 21 HAGEGMWVP 29

RESULT 9  
 ACFFAL  
 nicotinic acetylcholine receptor alpha-1 chain precursor - fruit fly (Drosophila mela  
 C:Species: Drosophila melanogaster  
 C:Date: 31-Dec-1991 #sequence\_revision 31-Dec-1991 #text\_change 22-Jun-1999  
 C:Accession: S00381; A38801  
 R:Bossy, B.; Ballivet, M.; Spierer, P.  
 EMBO J. 7, 611-618, 1988  
 A:Title: Conservation of neural nicotinic acetylcholine receptors from Drosophila to  
 A:Reference number: S00381; MUID:88283626  
 A:Accession: S00381  
 A:Molecule type: DNA  
 A:Residues: 1-567 <BOS>  
 A:Cross-references: GB:X07194; NID:g7575; PIDN:CAA30172.1; PID:g7576  
 A:Accession: A38801  
 A:Molecule type: mRNA  
 A:Residues: 1-567 <BO2>  
 A:Cross-references: EMBL:X07194; NID:g7575; PIDN:CAA30172.1; PID:g7576  
 A:Note: 538-Tyr was also found  
 C:Genetics:  
 A:Gene: FlyBase:nAcr4agr  
 A:Cross-references: FlyBase:FBgn0000036  
 A:Map position: 3R 96A  
 A:Introns: 64/3; 79/3; 116/2; 176/3; 330/2; 401/1; 491/1  
 C:Superfamily: acetylcholine receptor  
 C:Keywords: glycoprotein; ion channel; neurotransmitter receptor; postsynaptic membra  
 F:1-21/Domain: signal sequence #status predicted <SIG>  
 F:22-567/Product: nicotinic acetylcholine receptor alpha-like chain #status predicted  
 F:22-240/Domain: extracellular #status predicted <EXT>  
 F:240-264/Domain: transmembrane #status predicted <TM1>  
 F:272-290/Domain: transmembrane #status predicted <TM2>  
 F:306-335/Domain: transmembrane #status predicted <TM3>  
 F:326-513/Domain: intracellular #status predicted <INT>  
 F:514-532/Domain: transmembrane #status predicted <TM4>  
 F:45,233/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 F:149-163,222-223/Disulfide bonds: #status predicted

Query Match 66.7%; Score 36; DB 1; Length 567;  
 Best Local Similarity 66.7%; Pred. No. 49;  
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 HTGXGVWXP 11  
 ||| |||  
 Db 133 HTGVVWXP 141

RESULT 10  
B75615  
pyridoxamine kinase - Deinococcus radiodurans (strain R1)  
C:Species: Deinococcus radiodurans  
C>Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 31-Mar-2000  
C:Accession: B75615  
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;  
S.; Smith, H.O.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma  
Science 286, 1571-1577, 1999  
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
A:Reference number: A75250; MUID:20036896  
A:Accession: B75615  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-329 <WHI>  
A:Cross-references: GB:AF001862; GB:AF001862; NID:g6460468; PIDN:AAF12189.1; PID:g646048  
A:Experimental source: strain R1  
C:Genetics:  
A:Gene: DRA0184  
A:Map position: 2

Query Match 64.8%; Score 35; DB 2; Length 329;  
Best Local Similarity 71.4%; Pred. No. 44;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 HTGXGVW 9  
||| | | | |

Db 82 HTGYGAW 88

RESULT 11  
A83673  
hypothetical protein BH0185 [imported] - Bacillus halodurans (strain C-125)  
C:Species: Bacillus halodurans  
C>Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 15-Jun-2001  
C:Accession: A83673  
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira  
Nucleic Acids Res. 28, 4317-4331, 2000  
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and  
A:Reference number: A83650; MUID:20512582; PMID:11058132  
A:Accession: A83673  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-448 <STO>  
A:Cross-references: GB:AP001507; GB:BA000004; NID:g10172612; PIDN:BAR03904.1; GSPDB:GN00  
A:Experimental source: strain C-125  
C:Genetics:  
A:Gene: BH0185

Query Match 64.8%; Score 35; DB 2; Length 448;  
Best Local Similarity 55.6%; Pred. No. 60;  
Matches 5; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 HTGXGVWXP 11  
| | | | | | |

Db 220 HTGFGTWGP 228

RESULT 12  
T46589  
ropy-2 protein [imported] - Neurospora crassa  
C:Species: Neurospora crassa  
C>Date: 18-Feb-2000 #sequence\_revision 18-Feb-2000 #text\_change 04-Mar-2000  
C:Accession: T46589  
R:Vierula, J.P.; Mals, J.M.  
submitted to the EMBL Data Library, March 1995  
A:Description: Cloning and characterization of the ropy-2 gene from Neurospora crassa an  
A:Reference number: Z23094  
A:Accession: T46589

A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-710 <VIE>  
A:Cross-references: EMBL:U23425; PIDN:AA868457.1  
A:Experimental source: strain 74-OR23-1VA; clone cJH1  
C:Genetics:  
A:Gene: ropy-2  
A:Introns: 276/2  
C:Superfamily: Neurospora crassa ropy-2 protein

Query Match 64.8%; Score 35; DB 2; Length 710;  
Best Local Similarity 62.5%; Pred. No. 95;  
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 4 TGXGVWXP 11  
||| | | | |

Db 135 TGSGLWFP 142

RESULT 13  
B86687  
conserved hypothetical protein yejH [imported] - Lactococcus lactis subsp. lactis (st  
C:Species: Lactococcus lactis subsp. lactis  
C>Date: 23-Mar-2001 #sequence\_revision 23-Mar-2001 #text\_change 03-Aug-2001  
C:Accession: B86687  
R:Boilotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarme, K.; Weissbach, J.; Eh  
Genome Res. 11, 731-753, 2001  
A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis  
A:Reference number: A86625; MUID:21235186; PMID:11337171  
A:Accession: B86687  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-285 <STO>  
A:Cross-references: GB:AE005176; PID:g12723380; PIDN:AAK04596.1; GSPDB:GN00146  
A:Experimental source: strain ILL403  
C:Genetics:  
A:Gene: yejH

Query Match 63.0%; Score 34; DB 2; Length 285;  
Best Local Similarity 71.4%; Pred. No. 58;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 HTGXGVW 9  
||| | | | |

Db 272 HTGPGAW 278

RESULT 14  
S27600  
N-acetylglutamate synthase - Pseudomonas aeruginosa  
C:Species: Pseudomonas aeruginosa  
C>Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 15-Sep-2000  
C:Accession: S27600  
R:Dharmsthiti, S.; Krishnapillai, V.  
submitted to the EMBL Data Library, February 1991  
A:Description: Reduced sequence conservation in the argA genes of two closely related  
A:Reference number: S27600  
A:Accession: S27600  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-397 <DHA>  
A:Cross-references: EMBL:M38358; NID:g151035; PIDN:AAA73377.1; PID:g151036  
C:Superfamily: amino-acid acetyltransferase

Query Match 63.0%; Score 34; DB 2; Length 397;  
Best Local Similarity 62.5%; Pred. No. 81;  
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 TGXGVWXP 11  
||| | | | |



Db 265 TNGTWS 272

RESULT 15

B98354  
 hypothetical protein AGR\_L3555 [imported] - Agrobacterium tumefaciens (strain C58, Cere  
 C:Species: Agrobacterium tumefaciens  
 C>Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 11-Jan-2002  
 C:Accession: B98354  
 R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,  
 A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;  
 Science 294, 2323-2328, 2001  
 A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum  
 A:Reference number: A97359; PMID:11743194  
 A:Accession: B98354  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-414 <KUR>  
 A:Cross-references: GB:AE007870; PIDN:AAK90356.1; PID:gl5160397; GSPDB:GN00170  
 C:Genetics:  
 A:Gene: AGR\_L3555  
 A:Map position: linear chromosome

Query Match 63.0%; Score 34; DB 2; Length 414;  
 Best Local Similarity 71.4%; Pred. No. 85;  
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 HTGXGWN 9

Db 304 HPGTGW 310

Search completed: October 4, 2002, 10:26:18  
 Job time: 345 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 4, 2002, 10:22:14 ; Search time 20.58 Seconds  
(without alignments)  
16.616 Million cell updates/sec

Title: US-09-822-698A-28  
Perfect score: 54  
Sequence: 1 XXHTGXGVWXPXX 14

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA.\*  
1: /cgn2.6/prodata/2/1aa/5A\_COMB.pap.\*  
2: /cgn2.6/prodata/2/1aa/5B\_COMB.pap.\*  
3: /cgn2.6/prodata/2/1aa/6A\_COMB.pap.\*  
4: /cgn2.6/prodata/2/1aa/6B\_COMB.pap.\*  
5: /cgn2.6/prodata/2/1aa/PCTUS\_COMB.pap.\*  
6: /cgn2.6/prodata/2/1aa/backfiles1.pap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	38	70.4	656	1	US-08-444-005-15
2	38	70.4	656	4	US-09-069-023-28
3	34	63.0	336	2	US-08-997-080-156
4	34	63.0	336	2	US-08-997-362-156
5	34	63.0	336	4	US-09-095-855-156
6	32	59.3	120	6	5514582-36
7	32	59.3	573	1	US-08-200-512-2
8	31.5	58.3	774	4	US-09-346-237-8
9	31	57.4	128	1	US-08-478-039-73
10	31	57.4	128	1	US-08-478-039-102
11	31	57.4	128	1	US-08-478-349A-73
12	31	57.4	128	1	US-08-478-349A-102
13	31	57.4	262	1	US-08-392-828C-37
14	31	57.4	262	3	US-09-330-945-37
15	31	57.4	321	2	US-08-712-072C-3
16	31	57.4	365	3	US-08-155-005A-8
17	31	57.4	365	4	US-09-363-783-8
18	31	57.4	388	3	US-08-155-005A-6
19	31	57.4	388	4	US-09-363-783-6
20	31	57.4	390	3	US-08-155-005A-4
21	31	57.4	390	3	US-08-155-005A-17
22	31	57.4	390	4	US-09-363-783-4
23	31	57.4	390	4	US-09-363-783-17
24	31	57.4	523	3	US-08-651-999A-2
25	31	57.4	523	4	US-09-385-752-2
26	31	57.4	4302	3	US-08-658-136-5
27	31	57.4	4303	2	US-08-460-751-2

28	30	55.6	712	1	US-08-587-889-2	Sequence 2, Appli
29	30	55.6	712	2	US-08-980-060-5	Sequence 5, Appli
30	30	55.6	712	4	US-09-307-185-5	Sequence 5, Appli
31	30	55.6	712	5	PCT-US96-09193-2	Sequence 2, Appli
32	30	55.6	739	3	US-09-136-251-2	Sequence 2, Appli
33	30	55.6	984	4	US-09-287-354-2	Sequence 2, Appli
34	30	55.6	1182	4	US-09-287-354-6	Sequence 6, Appli
35	30	55.6	1189	4	US-09-287-354-3	Sequence 3, Appli
36	30	55.6	1189	4	US-09-287-354-4	Sequence 4, Appli
37	30	55.6	1207	4	US-09-287-354-5	Sequence 5, Appli
38	30	55.6	1865	1	US-08-588-985-2	Sequence 2, Appli
39	30	55.6	1865	1	US-08-571-988-2	Sequence 2, Appli
40	29.5	54.6	209	3	US-08-899-578-6	Sequence 2, Appli
41	29.5	54.6	587	3	US-08-899-578-2	Sequence 2, Appli
42	29	53.7	18	2	US-08-934-915-184	Sequence 184, App
43	29	53.7	20	2	US-08-934-915-179	Sequence 179, App
44	29	53.7	126	6	5514582-43	Patent No. 5514582
45	29	53.7	234	1	US-08-152-456A-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1  
US-08-444-005-15  
; Sequence 15, Application US/08444005  
; Patent No. 5674734  
; GENERAL INFORMATION:  
; APPLICANT: Leder, Philip  
; APPLICANT: Seed, Brian  
; APPLICANT: Stanger, Ben Z.  
; APPLICANT: Lee, Tae-Ho  
; APPLICANT: Kim, Emily  
; TITLE OF INVENTION: CELL DEATH PROTEIN  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 225 Franklin Street, Suite 3100  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/444,005  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Clark, Paul T.  
; REGISTRATION NUMBER: 30,164  
; REFERENCE/DOCKET NUMBER: 00383/026001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617/542-5070  
; TELEFAX: 617/542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 15:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 656 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: not relevant  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-444-005-15

Query Match 70.4%; Score 38; DB 1; Length 656;  
Best Local Similarity 75.08; Pred. No. 29;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 TGXGVWXP 11  
Dd 473 TGTGVWYP 480

```

RESULT      2
US-09-069-023-28
: Sequence 28, Application US/09069023A
: Patent No. 6348573
: GENERAL INFORMATION:
: APPLICANT: Nunez, Gabriel
: APPLICANT: Inohara, Naohiro
: APPLICANT: Koseki, Takeyoshi
: TITLE OF INVENTION: COMPOSITIONS AND METHODS
: FILE REFERENCE: UM-0333
: CURRENT APPLICATION NUMBER: US/09/069, 023A
: CURRENT FILING DATE: 1998-04-27
: NUMBER OF SEQ ID NOS: 38
: SOFTWARE: Patentin Ver. 2.0
: SEQ ID NO 28
: LENGTH: 656
: TYPE: PRT
: ORGANISM: Mus musculus
US-09-069-023-28

```

Query Match 70.4%; Score 38; DB 4; Length 656;  
Best Local Similarity 75.0%; Pred. No. 29;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 TGXGVWXP 11  
Db 473 TGTGVWYP 480

RESULT 3  
 US-08-997-060-156  
 : Sequence 156, Application US/08997080  
 : Patent No. 5968524  
 : GENERAL INFORMATION:  
 : APPLICANT: WATSON, JAMES D.  
 : APPLICANT: TAN, PAUL L.J.  
 : TITLE OF INVENTION: METHODS AND COMPOUNDS FOR THE TREATMENT OF IMMUNOLOGICALLY-  
 : NUMBER OF SEQUENCES: 194  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: Law Offices of Ann W. Speckman  
 : STREET: 2601 Elliott Avenue, Suite 4185  
 : CITY: Seattle  
 : STATE: WA  
 : COUNTRY: USA  
 : ZIP: 98121  
 : COMPUTER READABLE FORM:  
 : MEDIUM TYPE: diskette  
 : COMPUTER: IBM compatible  
 : OPERATING SYSTEM: DOS  
 : SOFTWARE: Fastseq for Windows Version 2.0  
 : CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: US/08/997,080  
 : FILING DATE:  
 : CLASSIFICATION:  
 : PRIOR APPLICATION DATA:  
 : APPLICATION NUMBER:  
 : FILING DATE:  
 : ATTORNEY/AGENT INFORMATION:  
 : NAME: Sleath, Janet  
 : REGISTRATION NUMBER: 37,007  
 : REFERENCE/DOCKET NUMBER: 11000.1007  
 : TELECOMMUNICATION INFORMATION:  
 : TELEPHONE: 206-269-0565  
 : TELEFAX: 206-269-0563  
 : TELEX:  
 : INFORMATION FOR SEQ ID NO: 156:

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 336 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-997-080-156

```

Query Match 63.0%; Score 34; DB 2; Length 336;  
Best Local Similarity 71.4%; Pred. No. 77;  
Matches 5: Conservative 0; Mismatches 2: Indels

Qy 5 GXGVWXP 11  
| | | | |  
Db 288 GSGVWAP 294

RESULT 4  
US-08-997-362-156  
; Sequence 156, Application US/08997362  
; Patent No. 5985287  
; GENERAL INFORMATION:  
; APPLICANT: Tan, Paul  
; APPLICANT: Hiyama, Jun  
; APPLICANT: Visser, Elizabeth  
; APPLICANT: Skinner, Margot  
; APPLICANT: Scott, Linda  
; APPLICANT: Prestidge, Ross  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR  
; TREATMENT AND DIAGNOSIS OF MYCOBACTERIAL INFECTIONS  
; NUMBER OF SEQUENCES: 194  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Law Offices of Ann W. Speckman  
; STREET: 2601 Elliott Avenue, Suite 4185  
; CITY: Seattle  
; STATE: WA  
; COUNTRY: USA  
; ZIP: 98121  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/997, 362

Query Match 63.0%; Score 34; DB 2; Length 336; ,  
 Best Local Similarity 71.4%; Pred. No. 77;  
 Matches 5: Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 GXGVWXP 11  
1 111 1  
Db 288 GSGVWAP 294

## RESULT 5

US-09-095-855-156  
; Sequence 156, Application US/09095855  
; Patent No. 6160093  
; GENERAL INFORMATION:  
; APPLICANT: Tan, Paul  
; APPLICANT: Visser, Elizabeth  
; APPLICANT: Skinner, Margot  
; APPLICANT: Prestidge, Ross  
; TITLE OF INVENTION: Compounds and Methods for  
; TITLE OF INVENTION: Treatment and Diagnosis of Mycobacterial Infections  
; NUMBER OF SEQUENCES: 208  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Law Offices of Ann W. Speckman  
; STREET: 2601 Elliott Avenue, Suite 4185  
; CITY: Seattle  
; STATE: WA  
; COUNTRY: USA  
; ZIP: 98121  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/095,855  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/705,347  
; FILING DATE: 29-AUG-1996  
; APPLICATION NUMBER: 08/873,970  
; FILING DATE: 12-JUN-1997  
; APPLICATION NUMBER: 08/997,362  
; FILING DATE: 23-DEC-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sleath, Janet  
; REGISTRATION NUMBER: 37,007  
; REFERENCE/DOCKET NUMBER: 11000.1002c3  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 206-269-0565  
; TELEFAX: 206-269-0563  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 156:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 336 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-09-095-855-156

Query Match 63.0%; Score 34; DB 4; Length 336;  
Best Local Similarity 71.4%; Pred. No. 77;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 GXGVWXP 11  
1 111 1  
Db 288 GSGVWAP 294

## RESULT 6

5514582-36  
; Patent No. 5514582  
; APPLICANT: CAPON, DANIEL J.; LASKY, LAURENCE A.  
; TITLE OF INVENTION: RECOMBINANT DNA ENCODING HYBRID  
; IMMUNOGLOBULINS  
; NUMBER OF SEQUENCES: 43

; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/185,670  
; FILING DATE: 21-JAN-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 986,931  
; FILING DATE: 08-DEC-1992  
; APPLICATION NUMBER: 808,122  
; FILING DATE: 16-DEC-1991  
; APPLICATION NUMBER: 440,625  
; FILING DATE: 22-NOV-1989  
; APPLICATION NUMBER: 315,015  
; FILING DATE: 23-FEB-1989  
; SEQ ID NO:36:  
; LENGTH: 120  
5514582-36

Query Match 59.3%; Score 32; DB 6; Length 120;  
Best Local Similarity 62.5%; Pred. No. 62;  
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 TGXGVWXP 11  
1 111 1  
Db 46 TSEGWTP 53

## RESULT 7

US-08-200-512-2  
; Sequence 2, Application US/08200512  
; Patent No. 5674718  
; GENERAL INFORMATION:  
; APPLICANT: BRIAND, MICHELLE  
; APPLICANT: PLESIAI, PATRICK  
; TITLE OF INVENTION: DNA SEGMENTS AND TRANS-  
; TITLE OF INVENTION: FORMED MICROORGANISMS COMPRISING THE DELTA'-  
; TITLE OF INVENTION: DEHYDROGENASE GENE, AND USES THEREOF  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BIERMAN & MUSERLIAN  
; STREET: 600 THIRD AVENUE  
; CITY: NEW YORK  
; STATE: NEW YORK  
; COUNTRY: USA  
; ZIP: 10016  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: FLOPPY DISK  
; COMPUTER: IBM PC COMPATIBLE  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WORDPERFECT 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/200,512  
; FILING DATE: 17-FEB-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/768,660  
; FILING DATE: 12-DEC-1991

; APPLICATION DATA:  
; APPLICATION NUMBER: PCT/FR91/00094  
; FILING DATE: 07-FEB-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: MUSERLIAN, CHARLES A.  
; REGISTRATION NUMBER: 19,683  
; REFERENCE/DOCKET NUMBER: 146.1116-CON  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-661-8000  
; TELEFAX: (212) 661-8002  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 573 AMINO ACIDS  
; TYPE: AMINO ACID  
; TOPOLOGY: LINEAR  
; MOLECULE TYPE:

```

; DESCRIPTION: PROTEIN
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: PSEUDOMONAS TESTOSTERONI
; FEATURE: (R, repeat if necessary)
; NAME/KEY: PROTEIN
; LOCATION: 1..573
; OTHER INFORMATION: /note = Delta
; OTHER INFORMATION: dehydrogenase protein of Pseudomonas
; OTHER INFORMATION: testosterone
;
US-08-200-512-2

```

```

Query Match          59.3%; Score 32; DB 1; Length 573;
Best Local Similarity 50.0%; Pred. No. 3e+02;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```

```

QY      4 TGXGVWXP 11
       :| | | |
Db      47 SGGIWP 54

```

```

RESULT      8
US-09-346-237-8
; Sequence 8, Application US/09346237A
; Patent No. 6265197
; GENERAL INFORMATION:
; APPLICANT: Bisgard-Frantzen, Henrik
; APPLICANT: Svendsen, Allan
; TITLE OF INVENTION: Starch Debranching Enzymes
; FILE REFERENCE: 5629.200-US
; CURRENT APPLICATION NUMBER: US/09/346.237A
; EARLIER FILING DATE: 1999-07-01
; EARLIER APPLICATION NUMBER: PA 1998 00868
; EARLIER FILING DATE: 1998-07-02
; EARLIER APPLICATION NUMBER: 60/094,353
; EARLIER FILING DATE: 1998-07-28
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 774
; TYPE: PRT
; ORGANISM: Favobacterium odoratum
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(774)
; OTHER INFORMATION: Isoamylase
;
US-09-346-237-8

```

```

Query Match          58.3%; Score 31.5; DB 4; Length 774;
Best Local Similarity 60.0%; Pred. No. 4.9e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

```

```

QY      3 HTG-XGVWXP 11
       :| | | |
Db      329 HTGEGGAWSP 338

```

```

RESULT      9
US-08-478-039-73
; Sequence 73, Application US/08478039
; Patent No. 5681722
; GENERAL INFORMATION:
; APPLICANT: Newman, Roland A.
; APPLICANT: Hanna, Nabil
; APPLICANT: Raab, Ronald W.
; TITLE OF INVENTION: Recombinant Antibodies for Human Therapy
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS
; STREET: 699 Prince St.

```

```

; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/478,039
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/379,072
; FILING DATE: 25-JAN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/912,292
; FILING DATE: 10-JUL-1992
; APPLICATION NUMBER: US 07/856,281
; FILING DATE: 23-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/735,064
; FILING DATE: 25-JUL-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Teskin Esq., Robin L.
; REGISTRATION NUMBER: 35,030
; REFERENCE/DOCKET NUMBER: 012712-160
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-836-6620
; TELEFAX: 703-836-2021
; INFORMATION FOR SEQ ID NO: 73:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 128 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Monkey
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT: VH3 clone 3-40
;
US-08-478-039-73

```

```

Query Match          57.4%; Score 31; DB 1; Length 128;
Best Local Similarity 62.5%; Pred. No. 99;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY      4 TGXGVWXP 11
       :| | | |
Db      113 TGSDVWGP 120

```

```

RESULT      10
US-08-478-039-102
; Sequence 102, Application US/08478039
; Patent No. 5681722
; GENERAL INFORMATION:
; APPLICANT: Newman, Roland A.
; APPLICANT: Hanna, Nabil
; APPLICANT: Raab, Ronald W.
; TITLE OF INVENTION: Recombinant Antibodies for Human Therapy
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS
; STREET: 699 Prince St.
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-1404
; COMPUTER READABLE FORM:

```

```
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/478,039
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/379,072
; FILING DATE: 25-JAN-1995
; APPLICATION NUMBER: US 07/912,292
; FILING DATE: 10-JUL-1992
; APPLICATION NUMBER: US 07/735,064
; FILING DATE: 25-JUL-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Teskin Esq., Robin L.
; REGISTRATION NUMBER: 35,030
; REFERENCE/DOCKET NUMBER: 012712-160
; TELEPHONE: 703-836-6620
; TELEFAX: 703-836-6620
; INFORMATION FOR SEQ ID NO: 102:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 128 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Monkey
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT: #40
US-08-478-039-102

Query Match 57.4%; Score 31; DB 1; Length 128;
Best Local Similarity 62.5%; Pred. No. 99;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 TGXGVWXP 11
DB 113 TGSVDWGP 120

RESULT 11
US-08-476-349A-73
; Sequence 73, Application US/08476349A
; Patent No. 5750105
; GENERAL INFORMATION:
; APPLICANT: Newman, Roland A.
; APPLICANT: Hanna, Nabil
; APPLICANT: Raab, Ronald W.
; TITLE OF INVENTION: Recombinant Antibodies for Human Therapy
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS
; STREET: 699 Prince St.
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
```

```
; APPLICATION NUMBER: US/08/476,349A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/379,072
; FILING DATE: 25-JAN-1995
; APPLICATION NUMBER: US 07/912,292
; FILING DATE: 10-JUL-1992
; APPLICATION NUMBER: US 07/856,281
; FILING DATE: 23-MAR-1992
; APPLICATION NUMBER: US 07/735,064
; FILING DATE: 25-JUL-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Teskin Esq., Robin L.
; REGISTRATION NUMBER: 35,030
; REFERENCE/DOCKET NUMBER: 012712-161
; TELEPHONE: 703-836-6620
; TELEFAX: 703-836-2021
; INFORMATION FOR SEQ ID NO: 73:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 128 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Monkey
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT: VH3 clone 3-40
US-08-476-349A-73

Query Match 57.4%; Score 31; DB 1; Length 128;
Best Local Similarity 62.5%; Pred. No. 99;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 TGXGVWXP 11
DB 113 TGSVDWGP 120

RESULT 12
US-08-476-349A-102
; Sequence 102, Application US/08476349A
; Patent No. 5750105
; GENERAL INFORMATION:
; APPLICANT: Newman, Roland A.
; APPLICANT: Hanna, Nabil
; APPLICANT: Raab, Ronald W.
; TITLE OF INVENTION: Recombinant Antibodies for Human Therapy
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS
; STREET: 699 Prince St.
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,349A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/379,072
```

;; FILING DATE: 25-JAN-1995  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/912,292  
;; FILING DATE: 10-JUL-1992  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/856,281  
;; FILING DATE: 23-MAR-1992  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/735,064  
;; FILING DATE: 25-JUL-1991  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Teskin Esq., Robin L.  
;; REGISTRATION NUMBER: 35,030  
;; REFERENCE/DOCKET NUMBER: 012712-161  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 703-836-6620  
;; TELEFAX: 703-836-2021  
;; INFORMATION FOR SEQ ID NO: 102:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 128 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: not relevant  
;; TOPOLOGY: not relevant  
;; MOLECULE TYPE: peptide  
;; ORIGINAL SOURCE:  
;; ORGANISM: Monkey  
;; POSITION IN GENOME:  
;; CHROMOSOME/SEGMENT: #40  
US-08-476-349A-102

Query Match 57.4%; Score 31; DB 1; Length 128;  
Best Local Similarity 62.5%; Pred. No. 99;  
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 TCXGWXP 11  
DB 113 TGSVDWMP 120

RESULT 13  
US-08-828C-37  
; Sequence 37, Application US/08392828C  
; Patent No. 5795962  
; GENERAL INFORMATION:  
; APPLICANT: IWANAGA, SADAOKI  
; APPLICANT: MUTA, TATSUSHI  
; APPLICANT: SEKI, NORIYAKI  
; APPLICANT: ODA, TOSHIO  
; TITLE OF INVENTION: NOVEL POLYPEPTIDE AND DNA ENCODING  
; NUMBER OF SEQUENCES: 39  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PATENT ADMINISTRATOR, TESTA, HURWITZ &  
; ADDRESS: THIBEAULT  
; STREET: 53 STATE STREET  
; CITY: BOSTON  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/392,828C  
; FILING DATE: 28-FEB-1995  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CAMPBELL, PAULA A  
; REGISTRATION NUMBER: 32,503  
; REFERENCE/DOCKET NUMBER: FJN-033

;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (617) 248-7000  
;; TELEFAX: (617) 248-7100  
;; INFORMATION FOR SEQ ID NO: 37:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 262 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; FEATURE:  
;; NAME/KEY: Protein  
;; LOCATION: 1..262  
;; OTHER INFORMATION: /note= "BGI A1 SEQUENCE (FIGURE 2)"  
US-08-392-828C-37

Query Match 57.4%; Score 31; DB 1; Length 262;  
Best Local Similarity 83.3%; Pred. No. 2e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 TCXGWM 9  
DB 107 TGDGWM 112

RESULT 14  
US-09-330-945-37  
; Sequence 37, Application US/09330945  
; Patent No. 6077946  
; GENERAL INFORMATION:  
; APPLICANT: IWANAGA, SADAOKI  
; APPLICANT: MUTA, TATSUSHI  
; APPLICANT: SEKI, NORIYAKI  
; APPLICANT: ODA, TOSHIO  
; TITLE OF INVENTION: DNA ENCODING HORSESHOE CRAB  
; TITLE OF INVENTION: AMEBOCYTE LYSATE FACTOR G SUBUNIT A  
; NUMBER OF SEQUENCES: 39  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PATENT ADMINISTRATOR, TESTA, HURWITZ &  
; ADDRESS: THIBEAULT, LLP  
; STREET: 125 HIGH STREET  
; CITY: BOSTON  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02110  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/330,945  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/119,995  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: FITCHER, EDMUND R  
; REGISTRATION NUMBER: 27,829  
; REFERENCE/DOCKET NUMBER: FJN-032DV  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 248-7000  
; TELEFAX: (617) 248-7100  
; INFORMATION FOR SEQ ID NO: 37:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 262 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FEATURE:



; NAME/KEY: Protein  
; LOCATION: 1..262  
; OTHER INFORMATION: /note= "BG1 A1 SEQUENCE (FIGURE 2)"  
US-09-330-945-37

Query Match 57.4%; Score 31; DB 3; Length 262;  
Best Local Similarity 83.3%; Pred. No. 2e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 TGXGVW 9  
|||

Db 107 TGDGVW 112

RESULT 15  
US-08-712-072C-3  
; Sequence 3, Application US/08712072C  
; Patent No. 5925541  
; GENERAL INFORMATION:  
; APPLICANT: Jack Goldstein, Alex Zhu and Lin Leng  
; TITLE OF INVENTION: ENDO-BETA-GALACTOSIDASE  
; NUMBER OF SEQUENCES: 13  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Amster, Rothstein & Ebenstein  
; STREET: 90 Park Avenue  
; CITY: New York  
; STATE: NY  
; COUNTRY: US  
; ZIP: 10016  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 INCH 1.44 Mb STORAGE DISKETTE  
; COMPUTER: IBM PC COMPATIBLE  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: ASCII  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/712,072C  
; FILING DATE: 11-SEP-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bogosian, Elizabeth A.  
; REGISTRATION NUMBER: 39,911  
; REFERENCE/DOCKET NUMBER: 63475/97  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 697-5995  
; TELEFAX: (212) 286-0854 or 286-0082  
; TELEX: TWX 710-581-4766  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 321 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE:  
; ORIGINAL SOURCE:  
; ORGANISM: el3b, Bacillus circulans  
US-08-712-072C-3

Query Match 57.4%; Score 31; DB 2; Length 321;  
Best Local Similarity 83.3%; Pred. No. 2.5e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 TGXGVW 9  
|||

Db 166 TGDGVW 171

Search completed: October 4, 2002, 10:22:15  
Job time: 183 sec

---

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 4, 2002, 10:29:42 ; Search time 50.76 Seconds  
(without alignments)  
30.635 Million cell updates/sec

Title: US-09-822-698A-28  
Perfect score: 14  
Sequence: 1 XXHTGXVWXPXXX 14

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 747574 seqs, 111073796 residues

Word size : 0

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : A\_Geneseq\_032802.\*

```

1: /SIDSI/gcgdata/hold-geneseg/geneseq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/hold-geneseg/geneseq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/hold-geneseg/geneseq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/hold-geneseg/geneseq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/hold-geneseg/geneseq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/hold-geneseg/geneseq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/hold-geneseg/geneseq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/hold-geneseg/geneseq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/hold-geneseg/geneseq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/hold-geneseg/geneseq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/hold-geneseg/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/hold-geneseg/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/hold-geneseg/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/hold-geneseg/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/hold-geneseg/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/hold-geneseg/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/hold-geneseg/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/hold-geneseg/geneseq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/hold-geneseg/geneseq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/hold-geneseg/geneseq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/hold-geneseg/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/hold-geneseg/geneseq-emb1/AA2001.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7	50.0	14	22	AAE12716
2	4	28.6	125	22	ABE10574
3	4	28.6	125	22	AAE92111
4	4	28.6	710	22	ABG20114
5	3	21.4	5	21	AAE84222
6	3	21.4	5	22	AAE97419
7	3	21.4	6	13	AAE20446
8	3	21.4	6	13	AAE20447
9	3	21.4	6	18	AAE22385
10	3	21.4	6	19	AAE75342
11	3	21.4	6	19	AAE75274

12	3	21.4	6	20	AAE49177	TATA box recognizi
13	3	21.4	6	20	AAE25854	Human secreted pro
14	3	21.4	6	21	ABE28300	Human secreted pep
15	3	21.4	7	5	AAE40089	Sequence at antige
16	3	21.4	7	19	AAE76696	Internal peptide 1
17	3	21.4	7	19	AAE73002	A. tubigenis LipA
18	3	21.4	7	21	AAE67571	Peptide exhibiting
19	3	21.4	8	10	AAE91585	Peptide fragment t
20	3	21.4	8	13	AAE28399	Prod. of PCR prime
21	3	21.4	8	13	AAE29930	Bone resorption ra
22	3	21.4	8	13	AAE29932	Hepatitis C virus
23	3	21.4	8	14	AAE36028	Hepatitis C virus
24	3	21.4	8	14	AAE36029	Hepatitis C virus
25	3	21.4	8	14	AAE36030	Hepatitis C virus
26	3	21.4	8	16	AAE78461	Collagen fragment
27	3	21.4	8	18	AAE30946	Birch pollen aller
28	3	21.4	8	18	AAE08264	Analogue of amino
29	3	21.4	8	18	AAE08290	Analogue of amino
30	3	21.4	8	18	AAE08291	Analogue of amino
31	3	21.4	8	18	AAE08292	Analogue of amino
32	3	21.4	8	18	AAE08296	Analogue of amino
33	3	21.4	8	18	AAE08226	Analogue of amino
34	3	21.4	8	18	AAE08228	Analogue of amino
35	3	21.4	8	18	AAE08230	Analogue of amino
36	3	21.4	8	18	AAE08268	Analogue of amino
37	3	21.4	8	18	AAE08218	Analogue of amino
38	3	21.4	8	18	AAE08219	Analogue of amino
39	3	21.4	8	18	AAE08222	Analogue of amino
40	3	21.4	8	18	AAE08223	Analogue of amino
41	3	21.4	8	20	AAE41855	Rheumatoid arthrit
42	3	21.4	8	20	AAE41930	Rheumatoid arthrit
43	3	21.4	8	20	AAE47032	Immunogenic peptid
44	3	21.4	8	20	AAE93870	Bifidobacterium bi
45	3	21.4	8	21	AAE84415	Amino acid sequenc

#### ALIGNMENTS

RESULT 1

ID AAE12716 standard; peptide; 14 AA.

AC AAE12716;

DT 04-JAN-2002 (first entry)

XX MUC1-specific binding member antigen binding domain.

XX Tumour-associated antigen mucin-1; MUC-1; adenocarcinoma; cytostatic; cancer; breast; ovary; lung; bladder; therapy.

OS Unidentified.

XX Key

XX Location/Qualifiers

FT Misc-difference 1

FT Misc-difference 2

FT Misc-difference 6

FT Misc-difference 10

FT Misc-difference 12

FT Misc-difference 13

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-difference 14

FT Misc-d

```

XX 30-MAR-2001; 2001WO-US10589.
PF 30-MAR-2000; 2000US-0538913.
XX (DYAX-) DYAX CORP.
XX Hoogenboom HRJM, Henderikx MPG;
XX WPI; 2001-626437/72.
XX Novel isolated tumor-associated antigen mucin-1-specific binding member
PT for diagnosing and treating cancer, comprises mucin-1 binding domain or
PT its portion for binding to an epitope of the protein core of mucin-1 -
XX Claim 1; Page 75; 126pp; English.
XX The invention relates to an isolated tumour-associated antigen mucin-1
CC (MUC-1)-specific binding member comprising an antigen binding domain
CC region having an antibody variable light (VL) or heavy (VH) region,
CC or a complementarity determining region (CDR) of VL or VH. MUC1-specific
CC binding member is useful for diagnosing cancer, preferably adenocarcinoma
CC The binding of MUC1-specific binding member to MUC1 is detected by a
CC magnetic resonance imaging, scintillation counting, and X-ray film.
CC MUC1-specific binding member is useful for treating cancer, preferably
CC adenocarcinoma, in an individual, where the cancer is present in tissue
CC of the breast, ovary, lung, or bladder of the individual. MUC1-specific
CC binding member is useful for diagnosing and imaging MUC1-expressing
CC cancer cells and tissues, for purifying or isolating non-glycosylated,
CC underglycosylated or cancer-associated forms of MUC1 or MUC1 epitope-
CC containing molecules, and for therapeutically or prophylactically
CC treating cancer. The present sequence is antigen binding domain of
CC MUC1-specific binding member.
XX SQ Sequence 14 AA;

Query Match 50.0%; Score 7; DB 22; Length 14;
Best Local Similarity 100.0%; Pred. NO. 0.0011;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 HTGCGVWXP 11
Db 3 htgcvwxp 11
|||||||

RESULT 2
ID ABB10574
ID ABB10574 standard; Protein; 125 AA.
XX AC ABB10574;
XX 14-JAN-2002 (first entry)
XX Human pancreatic cancer related polypeptide, SEQ ID NO: 223.
XX Human; cytostatic; antidiabetic; antiinflammatory; gastric; osteopathic;
XX antihormone; antitumor; thyroid-active; gene therapy; antitense therapy;
XX pancreatic cancer antigen inhibitor; pancreatic cancer; pancreatitis;
XX diabetes; endocrine disorder; acromegaly; hyperthyroidism;
XX gastrointestinal disorder; Crohn's disease; duodenal ulcer.
XX Homo sapiens.
XX WO200155206-A1.
XX 02-AUG-2001.
XX 17-JAN-2001; 2001WO-US01353.
XX 31-JAN-2000; 2000US-0179065.
XX 04-FEB-2000; 2000US-0180628.

```

```

PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226686.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 12-SEP-2000; 2000US-0232081.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 21-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 25-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.

```

PR 02-OCT-2000; 2000US-0237039.  
 PR 02-OCT-2000; 2000US-0237040.  
 PR 13-OCT-2000; 2000US-0239935.  
 PR 13-OCT-2000; 2000US-0239937.  
 PR 20-OCT-2000; 2000US-0240960.  
 PR 20-OCT-2000; 2000US-0241221.  
 PR 20-OCT-2000; 2000US-0241785.  
 PR 20-OCT-2000; 2000US-0241786.  
 PR 20-OCT-2000; 2000US-0241787.  
 PR 20-OCT-2000; 2000US-0241808.  
 PR 20-OCT-2000; 2000US-0241809.  
 PR 20-OCT-2000; 2000US-0241826.  
 PR 01-NOV-2000; 2000US-0244617.  
 PR 08-NOV-2000; 2000US-0246474.  
 PR 08-NOV-2000; 2000US-0246475.  
 PR 08-NOV-2000; 2000US-0246476.  
 PR 08-NOV-2000; 2000US-0246477.  
 PR 08-NOV-2000; 2000US-0246478.  
 PR 08-NOV-2000; 2000US-0246523.  
 PR 08-NOV-2000; 2000US-0246524.  
 PR 08-NOV-2000; 2000US-0246525.  
 PR 08-NOV-2000; 2000US-0246526.  
 PR 08-NOV-2000; 2000US-0246527.  
 PR 08-NOV-2000; 2000US-0246528.  
 PR 08-NOV-2000; 2000US-0246532.  
 PR 08-NOV-2000; 2000US-0246538.  
 PR 08-NOV-2000; 2000US-0246609.  
 PR 08-NOV-2000; 2000US-0246610.  
 PR 08-NOV-2000; 2000US-0246611.  
 PR 08-NOV-2000; 2000US-0246613.  
 PR 17-NOV-2000; 2000US-0246613.  
 PR 17-NOV-2000; 2000US-0249207.  
 PR 17-NOV-2000; 2000US-0249208.  
 PR 17-NOV-2000; 2000US-0249209.  
 PR 17-NOV-2000; 2000US-0249210.  
 PR 17-NOV-2000; 2000US-0249211.  
 PR 17-NOV-2000; 2000US-0249212.  
 PR 17-NOV-2000; 2000US-0249213.  
 PR 17-NOV-2000; 2000US-0249214.  
 PR 17-NOV-2000; 2000US-0249215.  
 PR 17-NOV-2000; 2000US-0249216.  
 PR 17-NOV-2000; 2000US-0249217.  
 PR 17-NOV-2000; 2000US-0249218.  
 PR 17-NOV-2000; 2000US-0249219.  
 PR 17-NOV-2000; 2000US-0249224.  
 PR 17-NOV-2000; 2000US-0249245.  
 PR 17-NOV-2000; 2000US-0249264.  
 PR 17-NOV-2000; 2000US-0249265.  
 PR 17-NOV-2000; 2000US-0249297.  
 PR 17-NOV-2000; 2000US-0249299.  
 PR 17-NOV-2000; 2000US-0249300.  
 PR 01-DEC-2000; 2000US-0250160.  
 PR 01-DEC-2000; 2000US-0250391.  
 PR 05-DEC-2000; 2000US-0251030.  
 PR 05-DEC-2000; 2000US-0251398.  
 PR 05-DEC-2000; 2000US-0256719.  
 PR 06-DEC-2000; 2000US-0251479.  
 PR 08-DEC-2000; 2000US-0251856.  
 PR 08-DEC-2000; 2000US-0251868.  
 PR 08-DEC-2000; 2000US-0251869.  
 PR 08-DEC-2000; 2000US-0251989.  
 PR 08-DEC-2000; 2000US-0251990.  
 PR 11-DEC-2000; 2000US-0254097.  
 PR 05-JAN-2001; 2001US-0259678.  
 (HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Barash SC, Ruben SM;  
 XX WPI; 2001-457717/49.  
 DR N-PSDB; ABA06904.

XX Isolated pancreatic cancer polypeptide for treating, preventing and/or  
 PT prognosing disorders related to the pancreas including pancreatic  
 PT cancers and also for testing and detection e.g. diagnosis -

PS Claim 11; SEQ ID NO 223; 537pp; English.  
 XX  
 CC The invention relates to an isolated polypeptide comprising an amino  
 CC acid sequence at least 90% identical to 188 amino acid sequences fully  
 CC defined in the specification and encoded by 188 cDNA clones fully  
 CC defined in the specification. The invention also relates to a fragment  
 CC having biological activity, a domain, an epitope, full length protein,  
 CC variant, allelic variant or a species homologue of the fully defined  
 CC sequence. The polynucleotide and polypeptide are useful for treating,  
 CC preventing and/or prognosing disorders related to the pancreas including  
 CC pancreatic cancer, pancreatitis, diabetes, endocrine disorders such as  
 CC acromegaly or hyperthyroidism, and gastrointestinal disorders such as  
 CC Crohn's disease and duodenal ulcers. The present sequence is a  
 CC pancreatic cancer-related polypeptide of the invention.  
 XX  
 SQ Sequence 125 AA;

Query Match 28.6%; Score 4; DB 22; Length 125;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
 Matches 5; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 4 TGXGV 8  
 Db 113 tgxgv 117

# RESULT 3

AAAM92111  
 ID AAAM92111 standard; Protein; 125 AA.

XX AC AAAM92111;

XX DT 06-NOV-2001 (first entry)

DE Human digestive system antigen SEQ ID NO: 1460.

KW Human; digestive system antigen; gene therapy; cancer; appendicitis;  
 KW ulcerative colitis; infection; Hirschsprung's disease; chronic colitis;  
 KW digestive system disorder; Meckel's diverticulum.

OS Homo sapiens.

XX FN WO200155314-A2.

XX PD 02-AUG-2001.

XX PF 17-JAN-2001; 2001WO-US01324.

XX PR 31-JAN-2000; 2000US-0179065.

PR 04-FEB-2000; 2000US-0180828.

PR 24-FEB-2000; 2000US-0184664.

PR 02-MAR-2000; 2000US-0186350.

PR 16-MAR-2000; 2000US-0189874.

PR 17-MAR-2000; 2000US-0190076.

PR 18-APR-2000; 2000US-0198123.

PR 19-MAY-2000; 2000US-0205515.

PR 07-JUN-2000; 2000US-0209467.

PR 28-JUN-2000; 2000US-0214886.

PR 30-JUN-2000; 2000US-0215135.

PR 07-JUL-2000; 2000US-0216647.

PR 07-JUL-2000; 2000US-0216880.

PR 11-JUL-2000; 2000US-0217487.

PR 11-JUL-2000; 2000US-0217496.

PR 14-JUL-2000; 2000US-0218290.

PR 26-JUL-2000; 2000US-0220963.

PR 26-JUL-2000; 2000US-0220964.

PR 14-AUG-2000; 2000US-0224518.

PR 14-AUG-2000; 2000US-0224519.

PR 14-AUG-2000; 2000US-0225213.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225266.

PR 14-AUG-2000; 2000US-0225267.

PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226688.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0232403.  
PR 14-SEP-2000; 2000US-0232404.  
PR 14-SEP-2000; 2000US-0232405.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0234984.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.

PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Barash SC, Ruben SM;  
XX WPI, 2001-502630/55.  
XX N-PSDB; AAK87884.  
XX  
XX Polynucleotides encoding digestive system antigens, useful for  
XX diagnosing, treating, preventing and/or prognosing disorders of the  
XX digestive system, particularly cancer and cancer metastases -  
PS Claim 11; SEQ ID NO 1460; 986pp; English.  
XX  
XX The present invention provides the protein and c-ding sequences of a  
XX number of human digestive system antigens. These can be used in the  
XX diagnosis, treatment and prevention of digestive system disorders,  
XX including cancer, Meckel's diverticulum, bacterial or parasitic  
XX infections, appendicitis, Hirschsprung's disease, chronic colitis or  
XX ulcerative colitis. The present sequence is a digestive system antigen of  
XX the invention.  
XX  
XX Sequence 125 AA;

Query Match 28.6%; Score 4; DB 22; Length 125;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 TCXGV 8  
DB 113 t9xgv 117

## RESULT 4

ABG20114  
 XX ABG20114 standard; Protein; 710 AA.  
 AC ABG20114;  
 XX DT 18-FEB-2002 (first entry)  
 XX DE Novel human diagnostic protein #20105.  
 XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder.  
 XX OS Homo sapiens.  
 XX PN WO200175067-A2.  
 XX PD 11-OCT-2001.  
 XX PF 30-MAR-2001; 2001WO-US08631.  
 XX PR 31-MAR-2000; 2000US-0540217.  
 XX PR 23-AUG-2000; 2000US-0649167.  
 XX PA (HYSE-) HYSEQ INC.  
 XX PI Drmanac RT, Liu C, Tang YT;  
 XX WPI; 2001-639362/73.  
 DR N-PSDB; AAS84301.  
 XX New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity -  
 XX Claim 20; SEQ ID No 50473; 103pp; English.  
 CC The invention relates to isolated polynucleotide (I) and  
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
 CC and gene mapping, and in recombinant production of (II). The  
 CC polynucleotides are also used in diagnostics as expressed sequence tags  
 CC for identifying expressed genes. (I) is useful in gene therapy techniques  
 CC to restore normal activity of (II) or to treat disease states involving  
 CC (II). (II) is useful for generating antibodies against it, detecting or  
 CC quantitating a polypeptide in tissue, as molecular weight markers and as  
 CC a food supplement. (II) and its binding partners are useful in medical  
 CC imaging of sites expressing (II). (I) and (II) are useful for treating  
 CC disorders involving aberrant protein expression or biological activity.  
 CC The polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG00010-ABG30377 represent novel human  
 CC diagnostic amino acid sequences of the invention.  
 CC Note: The sequence data for this patent did not appear in the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX Sequence 710 AA;

Query Match 28.6%; Score 4; DB 22; Length 710;  
 Best Local Similarity 100.0%; Pred. No. 8.5e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 HTGXG 7

Db 361 htgxg 365  
 |||||

## RESULT 5

AA84222  
 ID AAY84222 standard; peptide; 5 AA.  
 XX AC AAY84222;  
 XX DT 03-JUL-2000 (first entry)  
 XX DE Amino acid sequence of a peptide.  
 XX KW Modulator; sterol-regulated Site-1 protease; cholesterol;  
 KW sterol regulatory element binding protein; SREBP; lipid synthesis;  
 KW fatty acid biosynthesis; site-1 protease; protease inhibitor;  
 XX serum cholesterol; hypercholesterolemia; lipid metabolism.  
 XX OS Synthetic.  
 XX FH Key Location/Qualifiers  
 FT Modified-site 1 /note= "acetylated"  
 FT Misc-difference 6 /note= "attached to MCA"  
 FT WO200009677-A2.  
 XX PN 24-FEB-2000.  
 XX PD 13-AUG-1999; 99WO-US18544.  
 XX PR 14-AUG-1998; 98US-0096571.  
 XX PR 23-JUL-1999; 99US-0360237.  
 XX (TEXA ) UNIV TEXAS SYSTEM.  
 PA (BROW/) BROWN M S.  
 PA (CHEN/) CHENG D.  
 PA (ESPE/) ESPENSHADE P J.  
 PA (GOLD/) GOLDSTEIN J L.  
 PA (RAWS/) RAWSON R B.  
 PA (SAKA/) SAKAI J.  
 XX Brown MS, Cheng D, Espenshade PJ, Goldstein JL, Rawson RB;  
 PI Sakai J;  
 XX WPI; 2000-224327/19.  
 XX Novel assay for identifying modulators of sterol-regulated Site-1  
 PT protease useful for the treatment of hypercholesterolemia, involves  
 PT identifying an agent capable of down regulating Site-1 protease  
 PT activity -  
 XX Example 11; Page 107; 172pp; English.  
 XX The specification describes a method for identifying modulators of a  
 CC sterol-regulated Site-1 protease. Site-1 protease cleaves sterol  
 CC regulatory element binding proteins (SREBPs) in the endoplasmic  
 CC reticulum, initiating release from membranes and activating lipid  
 CC synthesis. The modulators therefore also modulate cholesterol and  
 CC fatty acid biosynthesis. The method comprises selecting an agent capable  
 CC of down regulating Site-1 protease and formulating a composition  
 CC comprising the agent. The site-1 protease inhibitors are useful for  
 CC treating a patient for elevated serum cholesterol. Diseases treated  
 CC include hypercholesterolemia and other lipid metabolism associated  
 CC conditions. The present sequence was used in the course of the  
 CC invention.  
 XX SQ Sequence 5 AA;

Query Match 21.4%; Score 3; DB 21; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 6.4e+05;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 HTG 5

|||

Db 3 htg 5

## RESULT 6

AAB97419  
ID AAB97419 standard; peptide; 5 AA.XX  
AC AAB97419;XX  
DT 27-JUL-2001 (first entry)XX  
DE Site-1 protease fusion protein plasmid linker peptide.XX  
KW Site-1 protease; SIP; sterol regulatory element binding protein; linker;  
KW SREBP; fatty acid synthesis; cholesterol homeostasis; triglyceride.XX  
OS Synthetic.XX  
PN WO200127138-A2.XX  
PD 19-APR-2001.XX  
PF 11-OCT-2000; 2000WO-US28428.XX  
PR 13-OCT-1999; 99US-0159236.XX  
PA (TULA-) TULARIK INC.XX  
PA (UYTE-) UNIV TEXAS SOUTHWESTERN MEDICAL CENT DAL.XX  
PI Jaen JC, Li L, Brown MS, Goldstein JL, Cheng D;XX  
DR WPI; 2001-328425/34.XX  
PT Peptide derivatives, useful as S1 protease inhibitors for modulating  
PT expression of genes regulated by SREBP transcription factors and for  
PT treating conditions associated with abnormal cholesterol or lipid  
PT homeostasis -XX  
PS Examples; Page 24; 41pp; English.XX  
CC The present invention describes a number of peptide derivatives which are  
CC capable of modulating Site-1 protease (SIP). SIP initiates the release of  
CC sterol regulatory element binding proteins (SREBPs), which control the  
CC synthesis of fatty acids and cholesterol. The peptide derivatives of the  
CC invention are, therefore, useful in the modulation of cholesterol  
CC homeostasis and in the treatment of conditions associated with abnormal  
CC levels of plasma cholesterol, lipoproteins or triglycerides. The present  
CC sequence is a peptide encoded by a linker sequence and used in plasmid  
CC construction in the exemplification of the invention.XX  
SQ Sequence 5 AA;

Query Match 21.4%; Score 3; DB 22; Length 5;

Best Local Similarity 100.0%; Pred. No. 6.4e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 HTG 5

Db 3 htg 5

## RESULT 7

AAR20446  
ID AAR20446 standard; Peptide; 6 AA.XX  
AC AAR20446;XX  
DT 14-APR-1992 (first entry)XX  
DE Anti-b-endorphin monoclonal antibody 3E7 high avidity peptide 40.

XX

KW Oligonucleotide library; receptor; Mab.

XX  
OS Synthetic.XX  
PN WO9119818-A.XX  
PD 26-DEC-1991.XX  
PF 19-JUN-1991; 91WO-US04384.XX  
PR 20-JUN-1990; 90US-0541108.XX  
PA (AFFV-) AFFYMAX TECHN NV.XX  
PI Dower WJ, Cwirla SE, Barrett RW;XX  
DR WPI; 1992-024434/03.XX  
PT Identifying peptide(s) which bind to a preselected receptor - by  
PT screening libraries expressing fusion proteins of bacteriophage  
PT coat proteinsXX  
PS Example; Fig 4; 68pp; English.XX  
CC The peptide was obtained from an N-terminal hexapeptide library which  
CC was screened with monoclonal antibody 3E7 by carrying out three  
CC rounds of panning, elution and amplification. The peptide sequence  
CC was deduced from the DNA. The antibody binds to B-endorphin and  
CC recognises the N-terminal portion of the protein (YGGF) which is  
CC present on most natural opioid peptides. The antibody also binds  
CC tightly to leu- and met-enkephalin (YGGFL and YGGWM). This method  
CC may be used to identify peptides which bind to selected receptors  
CC without any prior information on the structure of the expected  
CC ligand or receptor. See also AAR20405-R20456.XX  
SQ Sequence 6 AA;

Query Match 21.4%; Score 3; DB 13; Length 6;

Best Local Similarity 100.0%; Pred. No. 6.4e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 GYW 9

Db 2 gvw 4

## RESULT 8

AAR20447  
ID AAR20447 standard; Peptide; 6 AA.XX  
AC AAR20447;XX  
DT 14-APR-1992 (first entry)XX  
DE Anti-b-endorphin monoclonal antibody 3E7 high avidity peptide 41.XX  
KW Oligonucleotide library; receptor; Mab.XX  
OS Synthetic.XX  
PN WO9119818-A.XX  
PD 26-DEC-1991.XX  
PF 19-JUN-1991; 91WO-US04384.XX  
PR 20-JUN-1990; 90US-0541108.XX  
PA (AFFV-) AFFYMAX TECHN NV.XX  
PI Dower WJ, Cwirla SE, Barrett RW;

XX



DR WPI; 1992-024434/03.

XX Identifying peptide(s) which bind to a preselected receptor - by

PT screening libraries expressing fusion proteins of bacteriophage

PT coat proteins

XX Example; Fig 4; 68pp; English.

PS The peptide was obtained from an N-terminal hexapeptide library which

XX was screened with monoclonal antibody 3E7 by carrying out three

CC rounds of panning, elution and amplification. The peptide sequence

CC was deduced from the DNA. The antibody binds to B-endorphin and

CC recognises the N-terminal portion of the protein (YGGF) which is

CC present on most natural opioid peptides. The antibody also binds

CC tightly to leu- and met-enkephalin (YGGFL and YGGFM). This method

CC may be used to identify peptides which bind to selected receptors

CC without any prior information on the structure of the expected

CC ligand or receptor. See also AAR20405-R20456.

XX Sequence 6 AA;

SQ

Query Match 21.4%; Score 3; DB 13; Length 6;

Best Local Similarity 100.0%; Pred. No. 6.4e+05;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 GWV 9

Db |||

2 gvv 4

RESULT 9

AAW22385

ID AAW22385 standard; peptide; 6 AA.

XX AC AAW22385;

XX 20-JAN-1998 (first entry)

XX Gonadotropin releasing hormone conjugate linking group.

XX Gonadotropin releasing hormone; vaccine; immuno-sterilisation;

KW fertility control; steroid hormone stimulated tumour; endometriosis;

KW GnRH.

XX Synthetic.

OS Key Location/Qualifiers

XX Modified-site 5

FT /label= bala

XX WO9715316-A1.

XX 01-MAY-1997.

XX 23-OCT-1996; 96WO-US16950.

XX 13-FEB-1996; 96GB-0002951.

PR 27-OCT-1995; 95US-0005905.

XX (MERI ) MERCK & CO INC.

XX Lombardo VK, Marburg S, Tolman RL;

XX WPI; 1997-258759/23.

XX New gonadotropin releasing hormone conjugates - used as vaccines for

PT immuno-sterilisation of animals, fertility control or for treating

PT steroid hormone stimulated tumours or conditions

XX Claim 11; Page 68; 75pp; English.

XX A gonadotropin releasing hormone conjugate has been developed which has

CC the formula: 2-[L1-Am-(Y1(L2-X)-An)(q-Y2(L2-X)-Ap)(r), where A = an

CC amino acid selected from Gly, Ser, Thr, beta-Ala, and Ala, with the

CC proviso that at least one A is Ser or Thr; L1 = a linker optionally

CC attached to an internal marker; L2 = a linker; X = a gonadotrophin

CC releasing hormone (GnRH) having a free sulphydryl group; Y1, Y2 = Lys

CC or ornithine (Orn); Z = an immunogenic carrier protein; m = 0-3; n = 5

CC or 6; p = 0-1; q = 1 or 2; r = 1-3. The present sequence represents

CC a specifically claimed example of An. They are useful as vaccines

CC for use in immunostimulating animals, fertility control, and for

CC treatment of steroid hormone stimulated tumour or conditions such as

CC endometriosis. The conjugates employ a hydrophilic linear scaffold as a

CC means to increase loading of desirable hapten onto a carrier protein.

CC Compared to conventional GnRH conjugates, these provide more GnRH per

CC carrier molecule while utilising a minimal number of carrier molecule

CC attachment sites. The novel vaccine compositions provide a high titre of

CC anti-GnRH antibody and produce minimal injection site lesions.

XX Sequence 6 AA;

SQ

Query Match 21.4%; Score 3; DB 18; Length 6;

Best Local Similarity 100.0%; Pred. No. 6.4e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 TGXG 7

Db ||||

3 tgxg 6

RESULT 10

AAW75342

ID AAW75342 standard; peptide; 6 AA.

XX AC AAW75342;

XX 02-FEB-1999 (first entry)

XX Hexapeptide #7 binds anti-Ad5 fibre head MAB 1D6.3.

XX Cellular receptor; virus; immobilisation; monoclonal antibody; fibre;

KW viral surface protein; hexapeptide expression library; adenovirus;

KW major histocompatibility complex; MHC; fibronectin; gene therapy;

KW genetic disease; acquired immune deficiency syndrome; AIDS; cancer.

XX Synthetic.

OS FR2758821-A1.

XX 31-JUL-1998.

XX 30-JAN-1997; 97FR-0001005.

XX 30-JAN-1997; 97FR-0001005.

PR 09-SEP-1997; 97FR-0011166.

XX (CNRS ) CNRS CENT NAT RECH SCI.

XX Boulanger P, Hong SS, Karayan L;

XX WPI; 1998-416458/36.

XX Use of peptide(s) for facilitating or modulating attachment of

PT adenovirus to cells - useful for, e.g. treating or preventing

PT infection and improving uptake of gene therapy vectors

XX Example 2; Fig 1; 48pp; French.

XX The invention relates to methods for selecting and identifying a

CC cellular receptor for a virus, by immobilising, on a support, a

CC monoclonal antibody targeted to a viral surface protein that determines

CC attachment of the virus to the receptor. The immobilised antibody is

CC incubated with a hexapeptide expression library and peptides bound to the

CC immobilised antibody are eluted by competitive binding with recombinant

CC fragments of the viral surface protein. In a reverse method, the viral  
 CC surface protein is immobilised and incubated with the peptide library. In  
 CC this case, the bound peptides are eluted by competitive binding with the  
 CC monoclonal antibody. The hexapeptides AAW75336-W75345 represent  
 CC peptides isolated by biopanning the library with the immobilised  
 CC monoclonal antibody (MAB) 1D6.3 and eluted using an adenovirus type 5  
 CC fibre head protein (AAW75334).  
 CC The methods are used to identify peptides from MHC Class I and/or  
 CC type III fibronectin proteins that allow or facilitate attachment by  
 CC adenovirus (Ad) to host cells and/or entry into the cells, and to  
 CC identify ligands that modulate Ad infection mediated by these peptides,  
 CC e.g. to treat or prevent Ad infections or to facilitate infection by Ad  
 CC gene therapy vectors used to treat genetic diseases, acquired immune  
 CC deficiency syndrome or cancer.

XX Sequence 6 AA;

Query Match 21.4%; Score 3; DB 19; Length 6;  
 Best Local Similarity 100.0%; Pred. No. 6.4e+05;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 3 HTG 5  
 |||  
 Db 4 htg 6

RESULT 11

AAW75274  
 ID AAW75274 standard; peptide; 6 AA.

XX  
 AC AAW75274;

DT 02-FEB-1999 (first entry)

DE Hexapeptide #7 binds anti-Ad5 fibre head MAB 1D6.3.

XX Cellular receptor; virus; immobilisation; monoclonal antibody; fibre;  
 KW viral surface protein; hexapeptide expression library; adenovirus;  
 KW major histocompatibility complex; MHC; fibronectin; gene therapy;  
 KW genetic disease; acquired immune deficiency syndrome; AIDS; cancer.  
 XX Synthetic.

XX FR275885-A1.

PN 31-JUL-1998.

PD 30-JAN-1997; 97FR-0001005.

PF 30-JAN-1997; 97FR-0001005.

PR 10-JUL-1997; 97FR-0008796.

XX (CNRS ) CNRS CENT NAT RECH SCI.

XX Boulanger P, Hong SS, Karayan L;

XX WPI; 1998-416493/36.

XX Selection and identification of cellular receptors for viruses -  
 PT used to control attachment and entry of adenovirus into cells, e.g.  
 PT for treating infection or in gene therapy

XX Example 2; Fig 1; 43pp: French.

XX The invention relates to methods for selecting and identifying a  
 CC cellular receptor for a virus, by immobilising, on a support, a  
 CC monoclonal antibody targeted to a viral surface protein that determines  
 CC attachment of the virus to the receptor. The immobilised antibody is  
 CC incubated with a hexapeptide expression library and peptides bound to the  
 CC immobilised antibody are eluted by competitive binding with recombinant  
 CC fragments of the viral surface protein. In a reverse method, the viral  
 CC surface protein is immobilised and incubated with the peptide library. In

CC this case, the bound peptides are eluted by competitive binding with the  
 CC monoclonal antibody. The hexapeptides AAW75268-W75277 represent  
 CC peptides isolated by biopanning the library with the immobilised  
 CC monoclonal antibody (MAB) 1D6.3 and eluted using an adenovirus type 5  
 CC fibre head protein (AAW75266).  
 CC The methods are used to identify peptides from MHC Class I and/or  
 CC type III fibronectin proteins that allow or facilitate attachment by  
 CC adenovirus (Ad) to host cells and/or entry into the cells, and to  
 CC identify ligands that modulate Ad infection mediated by these peptides,  
 CC e.g. to treat or prevent Ad infections or to facilitate infection by Ad  
 CC gene therapy vectors used to treat genetic diseases, acquired immune  
 CC deficiency syndrome or cancer.

XX Sequence 6 AA;

Query Match 21.4%; Score 3; DB 19; Length 6;  
 Best Local Similarity 100.0%; Pred. No. 6.4e+05;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 3 HTG 5  
 |||  
 Db 4 htg 6

RESULT 12

AAW49177  
 ID AAY49177 standard; peptide; 6 AA.

XX  
 AC AAY49177;

DT 31-JAN-2000 (first entry)

DE TATA box recognizing zinc finger 2.

XX Zinc finger protein; DNA binding site; gene therapy; zif269; TATA box;  
 KW p53 binding site; nuclear receptor element; NRE.

XX Synthetic.

XX WO9948909-A2.

XX 30-SEP-1999.

XX 01-FEB-1999; 99WO-US02142.

XX 30-JAN-1998; 98US-0073223.

XX (GREI/) GREISMAN H A.

XX (PABO/) PABO C O.

XX Greisman HA, Pabo CO;

XX WPI; 1999-580414/49.

XX Creating new zinc finger proteins that bind to specific DNA sites,  
 PT useful tests and gene therapy -

XX Example 4; Fig 3A; 21pp: English.

XX The invention relates to a method for creating new zinc finger proteins  
 CC directed to specific DNA binding sites by adding one finger to the  
 CC protein at a time. The method is useful for creating new zinc finger  
 CC proteins that are directed to specific DNA binding sites. The proteins  
 CC may be useful for tests and gene therapy. The new method allows for the  
 CC interactions between each finger sub-unit of the protein in the design  
 CC of high-affinity proteins for specific target sites. Sequences  
 CC AAY49173-208 represent amino acid sequences of new zinc finger proteins  
 CC that recognize TATA box, p53 binding site and nuclear receptor element  
 CC (NRE).

XX Sequence 6 AA;

Query Match 21.4%; Score 3; DB 20; Length 6;  
 Best Local Similarity 100.0%; Pred. No. 6.4e+05;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 HTG 5  
 III  
 Db 2 htg 4

## RESULT 13

AAV25854  
 ID AAV25854 standard; Protein; 6 AA.

XX AC AAV25854;

XX DT 04-OCT-1999 (first entry)

XX DE Human secreted protein fragment encoded from gene 51.

XX KW Secreted protein; human; treatment; diagnosis; therapy; cancer; tumour;  
 KW neurodegenerative disorder; developmental abnormality; blood disorder;  
 KW fetal deficiency; blood disorder; leukemia; immune system; inflammation;  
 KW autoimmune disease; hepatic disease; renal disease; allergy; restenosis;  
 KW ischaemic shock; Alzheimer's disease; cognitive disorder; schizophrenia;  
 KW cardiovascular disorder; wound healing; stroke; arthritis; obesity;  
 KW asthma; sepsis; acne; psoriasis; transplant rejection; infection; AIDS;  
 KW metabolic disorder.

XX OS Homo sapiens.

XX PN WO9938881-A1.

XX PD 05-AUG-1999.

XX PF 27-JAN-1999; 99WO-US01621.

XX PR 30-JAN-1998; 98US-00731170.

XX PR 30-JAN-1998; 98US-00731159.

XX PR 30-JAN-1998; 98US-00731160.

XX PR 30-JAN-1998; 98US-00731161.

XX PR 30-JAN-1998; 98US-00731162.

XX PR 30-JAN-1998; 98US-00731164.

XX PR 30-JAN-1998; 98US-00731165.

XX PR 30-JAN-1998; 98US-00731167.

XX (HUMA-) HUMAN GENOME SCI INC.

XX PI Carter KC, Endress GA, Feng P, Ferrie AM, Florence C;

XX PI Florence KA, Janat F, Ni J, Rosen CA, Ruben SM;

XX PI Soppet DR, Young P, Yu G;

XX DR WPI; 1999-469315/39.

XX DR N-PSDB; AAZ00460.

XX PT New isolated human genes and the secreted polypeptides they encode

XX PT useful in, e.g. treatment of Alzheimer's

XX PS Disclosure; Page 368; 393pp; English.

XX CC This invention describes novel human genes (see AAZ00410-200477) and the  
 CC secreted proteins (see AAV25711-Y25778) and fragments (see  
 CC AAV25779-Y25907) they encode. The polynucleotides and their corresponding  
 CC secreted polypeptides are useful for preventing, treating or ameliorating  
 CC medical conditions e.g. by protein or gene therapy. Also pathological  
 CC conditions can be diagnosed by determining the amount of the new  
 CC polypeptides in a sample or by determining the presence of mutations in  
 CC the new polynucleotides. Specific uses are described for each of the 67  
 CC polynucleotides of the invention, based on which tissues they are most  
 CC highly expressed in, and include developing products for the diagnosis  
 CC or treatment of cancer, tumours, neurodegenerative disorders,  
 CC developmental abnormalities and fetal deficiencies, blood disorders,  
 CC leukemias, diseases of the immune system, autoimmune diseases, hepatic

CC and renal disease, inflammation, allergies, ischaemic shock, Alzheimer's  
 CC and cognitive disorders, schizophrenia, restenosis, cardiovascular  
 CC disorders, wound healing, stroke, arthritis, obesity, asthma, sepsis,  
 CC acne, psoriasis, transplant rejection, metabolic disorders, infections  
 CC and AIDS. The polypeptides are also useful for identifying their binding  
 CC partners.

XX SQ Sequence 6 AA;

Query Match 21.4%; Score 3; DB 20; Length 6;

Best Local Similarity 100.0%; Pred. No. 6.4e+05;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 GW 9

III

Db 1 gw 3

## RESULT 14

AAAB28300

ID AAB28300 standard; Peptide; 6 AA.

XX AC AAB28300;

XX DT 14-FEB-2001 (first entry)

XX DE Human secreted peptide #17 encoded by cDNA#23.

XX KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;  
 KW anti-allergic; hepatotropic; antidiabetic; antineoplastic; antitumor;  
 KW vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;  
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;  
 KW neurological disease; infection; human; secreted protein.

XX OS Homo sapiens.

XX PN WO200058355-A1.

XX PD 05-OCT-2000.

XX PF 22-MAR-2000; 2000WO-US07527.

XX PR 26-MAR-1999; 99US-0126501.

XX PR 22-DEC-1999; 99US-0171551.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Rosen CA, Ruben SM, Komatsoulis G;

XX DR WPI; 2000-602358/57.

XX DR N-PSDB; AAC67653.

XX PT Nucleic acid molecules encoding human secreted proteins, used in  
 preventing, treating or ameliorating a disorder, e.g. Alzheimer's and

XX PT Parkinson's diseases and cancers -

XX PS Claim 11; Page 346; 367pp; English.

XX CC Sequences AAB28278-B28326 represent the amino acid sequences of 50  
 CC human secreted proteins encoded by the genes AAB2831-C67680. The genes  
 CC and proteins are useful for preventing, ameliorating or treating medical  
 CC conditions, e.g. by protein or gene therapy. The genes are isolated from  
 CC a range of human tissues disclosed in the specification. The nucleic  
 CC acids, proteins, antibodies and (ant)agonists are useful in the  
 CC diagnosis, treatment and prevention of: (a) cancer, e.g. breast and  
 CC ovarian cancer, and other cancers of the adrenal gland, bone, bone  
 CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;  
 CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune  
 CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus,  
 CC Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative  
 CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;  
 CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and

Search completed: October 4, 2002, 10:29:42  
Job time: 234 sec

CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal  
CC and parasitic infections.  
XX  
SQ Sequence 6 AA;

Query Match 21.4%; Score 3; DB 21; Length 6;  
Best Local Similarity 100.0%; Pred. No. 6.4e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 HTG 5  
   |  
   |  
   |  
Db 3 htg 5

RESULT 15  
AAP40089  
ID AAP40089 standard; peptide; 7 AA.  
XX  
AC AAP40089;  
XX  
DT 01-FEB-1992 (first entry)  
XX  
DE Sequence at antigenic site of a VP1 capsid protein of foot and  
DE mouth disease virus (FMDV) selected from serotypes A24 Cruzeiro,  
DE C3 Indaial or O1 BFS.  
XX  
KW Foot and mouth disease; vaccine; treatment; prophylaxis; immunogen;  
KW capsid protein; VP1.  
XX  
OS Foot and mouth disease virus.  
XX  
PN EPI05481-A.  
XX  
PD 18-APR-1984.  
XX  
PF 29-SEP-1983; 83EP-0109769.  
XX  
PR 03-DEC-1982; 82GB-0034638.  
PR 30-SEP-1982; 82GB-0027860.  
XX  
PA (WELL ) WELLCOME FOUNDATION LTD.  
XX  
PI Boothroyd JC, Makoff AJ;  
XX  
DR WPI; 1984-102736/17.  
XX  
PT Foot and mouth disease virus antigens - comprising peptide(s)  
PT with aminoacid sequence of viral protein antigenic sites  
XX  
PS Claim 5b; Page 16; 24pp; English.  
XX  
CC The inventors claim peptides which have the AA sequence of the  
CC antigenic site of a VP1 protein selected from serotypes A24 Cruzeiro,  
CC C3 Indaial or O2 BFS, or fragments or derivs. of such sequences.  
CC Also claimed are immunogens comprising the peptides linked to a  
CC carrier, vaccines contg. the immunogens or peptide and DNA coding  
CC for the peptides. The vaccines may be used for prophylaxis of foot  
CC and mouth disease.  
XX  
SQ Sequence 7 AA;

Query Match 21.4%; Score 3; DB 5; Length 7;  
Best Local Similarity 100.0%; Pred. No. 6.4e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 HTG 5  
   |  
   |  
   |  
Db 4 htg 6





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 4, 2002, 10:30:09 ; Search time 20.62 Seconds  
(without alignments)  
16.584 Million cell updates/sec

Title: US-09-822-698A-28  
Perfect score: 14  
Sequence: 1 XXHTGXVWXPXXX 14

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 231628 seqs, 24425594 residues

Word size : 0

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents\_AA.\*  
1: /cgn2.6/ptodata/2/iaa/5A\_COMB.pep.\*  
2: /cgn2.6/ptodata/2/iaa/5B\_COMB.pep.\*  
3: /cgn2.6/ptodata/2/iaa/6A\_COMB.pep.\*  
4: /cgn2.6/ptodata/2/iaa/6B\_COMB.pep.\*  
5: /cgn2.6/ptodata/2/iaa/PCITUS\_COMB.pep.\*  
6: /cgn2.6/ptodata/2/iaa/backfiles.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3	21.4	7	4	US-09-169-015-8
2	3	21.4	8	1	US-07-872-644-19
3	3	21.4	8	1	US-08-297-494-19
4	3	21.4	8	1	US-08-297-510-19
5	3	21.4	8	1	US-08-446-206B-1
6	3	21.4	8	1	US-08-446-206B-6
7	3	21.4	8	1	US-08-446-206B-9
8	3	21.4	8	1	US-08-446-206B-10
9	3	21.4	8	1	US-08-446-206B-14
10	3	21.4	8	1	US-08-446-206B-15
11	3	21.4	8	1	US-08-446-206B-21
12	3	21.4	8	1	US-08-446-206B-22
13	3	21.4	8	1	US-08-446-206B-28
14	3	21.4	8	1	US-08-446-206B-29
15	3	21.4	8	1	US-08-446-206B-30
16	3	21.4	8	1	US-08-479-532-19
17	3	21.4	8	1	US-08-455-526-19
18	3	21.4	8	1	US-08-455-525-19
19	3	21.4	8	2	US-08-807-030-26
20	3	21.4	8	2	US-08-807-030-34
21	3	21.4	8	2	US-08-807-030-58
22	3	21.4	8	2	US-08-807-030-59
23	3	21.4	8	2	US-08-807-030-60
24	3	21.4	8	2	US-08-807-030-61
25	3	21.4	8	2	US-08-807-030-62
26	3	21.4	8	2	US-08-807-030-63
27	3	21.4	8	2	US-08-807-030-64

```

28 3 21.4 8 2 US-08-807-030-65 Sequence 65, Appl
29 3 21.4 8 2 US-08-807-030-66 Sequence 66, Appl
30 3 21.4 8 2 US-08-807-030-67 Sequence 67, Appl
31 3 21.4 8 2 US-08-807-030-68 Sequence 68, Appl
32 3 21.4 8 2 US-08-807-030-69 Sequence 69, Appl
33 3 21.4 8 2 US-08-807-030-70 Sequence 70, Appl
34 3 21.4 8 2 US-08-807-030-71 Sequence 71, Appl
35 3 21.4 8 2 US-08-807-030-72 Sequence 72, Appl
36 3 21.4 8 2 US-08-807-030-73 Sequence 73, Appl
37 3 21.4 8 2 US-08-807-030-74 Sequence 74, Appl
38 3 21.4 8 3 US-09-139-491-19 Sequence 19, Appl
39 3 21.4 8 4 US-08-444-818-512 Sequence 512, App
40 3 21.4 8 4 US-08-444-818-513 Sequence 513, App
41 3 21.4 8 4 US-08-444-818-514 Sequence 514, App
42 3 21.4 8 4 US-08-637-732A-34 Sequence 34, Appl
43 3 21.4 8 4 US-08-637-732A-36 Sequence 36, Appl
44 3 21.4 8 5 PCT-US92-03222-19 Sequence 19, Appl
45 3 21.4 8 6 5472884-2 Patent No. 5472884

```

#### ALIGNMENTS

```

RESULT 1
US-09-169-015-8
; Sequence 8, Application US/09169015
; Patent No. 6180343
; GENERAL INFORMATION:
; APPLICANT: Anderson, David
; TITLE OF INVENTION: Green Fluorescent Protein Fusions with Random Peptides
; FILE REFERENCE: A66900/DJB/RMS/SJR
; CURRENT APPLICATION NUMBER: US/09/169,015
; CURRENT FILING DATE: 1998-10-08
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 8
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-09-169-015-8

```

```

Query Match 21.4%; Score 3; DB 4; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 3 HTG 5
Db 5 HTG 7

```

```

RESULT 2
US-07-872-644-19
; Sequence 19, Application US/07872644
; Patent No. 5389527
; GENERAL INFORMATION:
; APPLICANT: Beavo, Joseph A.
; APPLICANT: Bentley, Kelley
; APPLICANT: Charbonneau, Harry
; APPLICANT: Sonnenburg, William K.
; TITLE OF INVENTION: DNA Encoding Mammalian
; TITLE OF INVENTION: Phosphodiesterases
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Bicknell
; STREET: Two First National Plaza, 20 South Clark
; STREET: Street
; CITY: Chicago
; STATE: Illinois

```

COUNTRY: USA  
ZIP: 60603  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/872,644  
FILING DATE: 19920420  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/688,356  
FILING DATE: 04-APR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 5389527and, Greta E.  
REGISTRATION NUMBER: 35,302  
REFERENCE/DOCKET NUMBER: 27866/30822  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 346-5750  
TELEFAX: (312) 984-9740  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-297-494-19

Query Match 21.4%; Score 3; DB 1; Length 8;  
Best Local Similarity 100.0%; Pred. No. 1.7e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 HTG 5  
DB 6 HTG 8

RESULT 4  
US-08-297-510-19  
Sequence 19, Application US/08297510  
Patent No. 5602019  
GENERAL INFORMATION:  
APPLICANT: Beavo, Joseph A.  
APPLICANT: Bentley, Kelley  
APPLICANT: Charbonneau, Harry  
APPLICANT: Sonnenburg, William K.  
TITLE OF INVENTION: DNA Encoding Mammalian  
NUMBER OF SEQUENCES: 58  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
STREET: Two First National Plaza, 20 South Clark  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60603  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.2  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/297,510  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/688,356  
FILING DATE: 04-APR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 5602019and, Greta E.  
REGISTRATION NUMBER: 35,302  
REFERENCE/DOCKET NUMBER: 27866/30822  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 346-5750  
TELEFAX: (312) 984-9740  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide

COUNTRY: USA  
ZIP: 60603  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/872,644  
FILING DATE: 19920420  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/688,356  
FILING DATE: 04-APR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 5389527and, Greta E.  
REGISTRATION NUMBER: 35,302  
REFERENCE/DOCKET NUMBER: 27866/30822  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 346-5750  
TELEFAX: (312) 984-9740  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-07-872-644-19

Query Match 21.4%; Score 3; DB 1; Length 8;  
Best Local Similarity 100.0%; Pred. No. 1.7e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 HTG 5  
DB 6 HTG 8

RESULT 3  
US-08-297-494-19  
Sequence 19, Application US/08297494  
Patent No. 5580771  
GENERAL INFORMATION:  
APPLICANT: Beavo, Joseph A.  
APPLICANT: Bentley, Kelley  
APPLICANT: Charbonneau, Harry  
APPLICANT: Sonnenburg, William K.  
TITLE OF INVENTION: DNA Encoding Mammalian  
NUMBER OF SEQUENCES: 58  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
STREET: Two First National Plaza, 20 South Clark  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60603  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/297,494  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/688,356  
FILING DATE: 04-APR-1991



US-08-297-510-19

Query Match 21.4%; Score 3; DB 1; Length 8;  
Best Local Similarity 100.0%; Pred.No. 1.7e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels

QY	3	HTG	5
Dd	6	HTG	8

```

RESULT      5
US-08-446-206B-1
; Sequence 1, Application US/08446206B
; Patent No. 5750647
; GENERAL INFORMATION:
; APPLICANT: Eyre, David R
; APPLICANT: Clemens, J Daniel
; APPLICANT: Ochs, Vincent W
; TITLE OF INVENTION: Synthetic Peptide Analogs of NTx
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Christensen, O'Connor, Johnson & Kindness PLLC

```

Query Match 21.4%; Score 3; DB 1; Length 8;

Best Local Similarity 100.0%; Pred. No. 1.7e+05;  
Matches 4; Conservative 0; Mismatches 0;

QY	5	GXGV	8
Dd	4	GXGV	7

```

: RESULT 6
: US-08-446-206B-6
: Sequence 6, Application US/08446206B
: Patent No. 5730647
: GENERAL INFORMATION:
: APPLICANT: Eyre, David R
: APPLICANT: Clemens, J Daniel
: APPLICANT: Ochs, Vincent W
: TITLE OF INVENTION: Synthetic Peptide Analogs of NTX
: NUMBER OF SEQUENCES: 31
: CORRESPONDENCE ADDRESS:
:

```

Query Match 21.4%; Score 3; DB 1; Length 8;  
Best Local Similarity 100.0%; Pred. No. 1.7e+05;  
Matches 4: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	5	GXGV	8
Db	4	GXGV	7

## 7 RESULT

US-08-446-206B-9  
; Sequence 9, Application US/08446206B  
; Patent No. 5750647  
; GENERAL INFORMATION:  
; APPLICANT: Eyre, David R  
; APPLICANT: Clemens, J Daniel  
; APPLICANT: Ochs, Vincent W  
; TITLE OF INVENTION: Synthetic Peptide Analogs of NTx  
; NUMBER OF SEQUENCES: 31  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Christensen, O'Connor, Johnson & Kindness PLLC  
; STREET: 1420 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: U.S.A.  
; ZIP: WA 98101  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/446,206B  
; FILING DATE: 19-MAY-1995  
; CLASSIFICATION: 530  
; PRIOR APPLICATION NUMBER:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Shelton, Dennis K  
; REGISTRATION NUMBER: 26,997  
; REFERENCE/DOCKET NUMBER: WROS18360  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 206 682 8100  
; TELEFAX: 206 224 0779  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 8 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FRAGMENT TYPE: N-terminal  
; FEATURE:  
; LOCATION: 1  
; OTHER INFORMATION: Xaa is glutamine or pyrrolidone  
; OTHER INFORMATION: carboxylic acid.  
; FEATURE:  
; LOCATION: 5  
; OTHER INFORMATION: Xaa is an amino acid with either a  
; OTHER INFORMATION: negatively charged R group, or a polar but uncharged  
; OTHER INFORMATION: R group, or a nonpolar (hydrophobic) R group.  
US-08-446-206B-9

Query Match 21.4%; Score 3; DB 1; Length 8;  
Best Local Similarity 100.0%; Pred. No. 1.7e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 GXGV 8  
||||  
Db 4 GXGV 7

RESULT 8  
US-08-446-206B-10  
; Sequence 10, Application US/08446206B  
; Patent No. 5750647  
; GENERAL INFORMATION:  
; APPLICANT: Eyre, David R  
; APPLICANT: Clemens, J Daniel  
; APPLICANT: Ochs, Vincent W

; TITLE OF INVENTION: Synthetic Peptide Analogs of NTx  
; NUMBER OF SEQUENCES: 31  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Christensen, O'Connor, Johnson & Kindness PLLC  
; STREET: 1420 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: U.S.A.  
; ZIP: WA 98101  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/446,206B  
; FILING DATE: 19-MAY-1995  
; CLASSIFICATION: 530  
; PRIOR APPLICATION NUMBER:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Shelton, Dennis K  
; REGISTRATION NUMBER: 26,997  
; REFERENCE/DOCKET NUMBER: WROS18360  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 206 682 8100  
; TELEFAX: 206 224 0779  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 8 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FRAGMENT TYPE: N-terminal  
; FEATURE:  
; LOCATION: 1  
; OTHER INFORMATION: Xaa is glutamine or pyrrolidone  
; OTHER INFORMATION: carboxylic acid.  
; FEATURE:  
; LOCATION: 5  
; OTHER INFORMATION: Xaa is an amino acid with either a  
; OTHER INFORMATION: negatively charged R group, or a polar but uncharged  
; OTHER INFORMATION: R group, or a nonpolar (hydrophobic) R group.  
US-08-446-206B-10

Query Match 21.4%; Score 3; DB 1; Length 8;  
Best Local Similarity 100.0%; Pred. No. 1.7e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 GXGV 8  
||||  
Db 4 GXGV 7

RESULT 9  
US-08-446-206B-14  
; Sequence 14, Application US/08446206B  
; Patent No. 5750647  
; GENERAL INFORMATION:  
; APPLICANT: Eyre, David R  
; APPLICANT: Clemens, J Daniel  
; APPLICANT: Ochs, Vincent W  
; TITLE OF INVENTION: Synthetic Peptide Analogs of NTx  
; NUMBER OF SEQUENCES: 31  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Christensen, O'Connor, Johnson & Kindness PLLC  
; STREET: 1420 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington

COUNTRY: U.S.A.  
ZIP: WA 98101  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/446,206B  
FILING DATE: 19-MAY-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Shelton, Dennis K  
REGISTRATION NUMBER: 26,997  
REFERENCE/DOCKET NUMBER: WROSL8360  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 206 682 8100  
TELEFAX: 206 224 0779  
TELEX:  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: N-terminal  
FEATURE:  
LOCATION: 1  
OTHER INFORMATION: Xaa is glutamine or pyrrolidone  
OTHER INFORMATION: carboxylic acid.  
FEATURE:  
LOCATION: 5  
OTHER INFORMATION: Xaa is an amino acid with either a  
OTHER INFORMATION: negatively charged R group, or a polar but uncharged  
OTHER INFORMATION: R group, or a nonpolar (hydrophobic) R group.

US-08-446-206B-14

Query Match 21.4%; Score 3; DB 1; Length 8;  
Best Local Similarity 100.0%; Pred. No. 1.7e-05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 GXGV 8  
||||  
Db 4 GXGV 7

RESULT 10  
US-08-446-206B-15  
Sequence 15, Application US/08446206B  
Patent No. 5750647  
GENERAL INFORMATION:  
APPLICANT: Eyre, David R  
APPLICANT: Clemens, J Daniel  
APPLICANT: Ochs, Vincent W  
TITLE OF INVENTION: Synthetic Peptide Analogs of NTx  
NUMBER OF SEQUENCES: 31  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Christensen, O'Connor, Johnson & Kindness PLLC  
STREET: 1420 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: U.S.A.  
ZIP: WA 98101  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/446,206B  
FILING DATE: 19-MAY-1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Shelton, Dennis K  
REGISTRATION NUMBER: 26,997  
REFERENCE/DOCKET NUMBER: WROSL8360  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 206 682 8100  
TELEFAX: 206 224 0779  
TELEX:  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: N-terminal  
FEATURE:  
LOCATION: 1  
OTHER INFORMATION: Xaa is glutamine or pyrrolidone  
OTHER INFORMATION: carboxylic acid.  
FEATURE:  
LOCATION: 5  
OTHER INFORMATION: Xaa is alpha-aminoadipic acid.  
US-08-446-206B-15

Query Match 21.4%; Score 3; DB 1; Length 8;  
Best Local Similarity 100.0%; Pred. No. 1.7e-05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 GXGV 8  
||||  
Db 4 GXGV 7

RESULT 11  
US-08-446-206B-21  
Sequence 21, Application US/08446206B  
Patent No. 5750647  
GENERAL INFORMATION:  
APPLICANT: Eyre, David R  
APPLICANT: Clemens, J Daniel  
APPLICANT: Ochs, Vincent W  
TITLE OF INVENTION: Synthetic Peptide Analogs of NTx  
NUMBER OF SEQUENCES: 31  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Christensen, O'Connor, Johnson & Kindness PLLC  
STREET: 1420 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: U.S.A.  
ZIP: WA 98101  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/446,206B  
FILING DATE: 19-MAY-1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Shelton, Dennis K

REGISTRATION NUMBER: 26,997  
REFERENCE/DOCKET NUMBER: WROS18360  
TELEPHONE: 206 682 8100  
TELEFAX: 206 224 0779  
TELEX:

INFORMATION FOR SEQ ID NO: 21:

SEQUENCE CHARACTERISTICS:

LENGTH: 8 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: peptide

FRAGMENT TYPE: N-terminal

FEATURE:

LOCATION: 1

OTHER INFORMATION: Xaa is glutamine or pyrrolidone

OTHER INFORMATION: carboxylic acid.

FEATURE:

LOCATION: 5

OTHER INFORMATION: Xaa is pyridinylalanine.

US-08-446-206B-21

Query Match

Best Local Similarity 21.4%; Score 3; DB 1; Length 8;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 GXGV 8

DB 4 GXGV 7

RESULT 12

US-08-446-206B-22

Sequence 22, Application US/08446206B

Patent No. 5750647

GENERAL INFORMATION:

APPLICANT: Eyre, David R

APPLICANT: Clemens, J Daniel

APPLICANT: Ochs, Vincent W

TITLE OF INVENTION: Synthetic Peptide Analogs of NTX

NUMBER OF SEQUENCES: 31

CORRESPONDENCE ADDRESS:

ADDRESSEE: Christensen, O'Connor, Johnson & Kindness PLLC

STREET: 1420 Fifth Avenue

CITY: Seattle

STATE: Washington

COUNTRY: U.S.A.

ZIP: WA 98101

COMPUTER READABLE FORM:

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/446,206B

FILING DATE: 19-MAY-1995

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Shelton, Dennis K

REGISTRATION NUMBER: 26,997

REFERENCE/DOCKET NUMBER: WROS18360

TELECOMMUNICATION INFORMATION:

TELEPHONE: 206 682 8100

TELEFAX: 206 224 0779

TELEX:

INFORMATION FOR SEQ ID NO: 22:

SEQUENCE CHARACTERISTICS:

LENGTH: 8 amino acids

TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: N-terminal  
FEATURE:  
LOCATION: 1  
OTHER INFORMATION: Xaa is glutamine or pyrrolidone  
OTHER INFORMATION: carboxylic acid.  
FEATURE:  
LOCATION: 5  
OTHER INFORMATION: Xaa is an alpha-amino acid wherein the  
OTHER INFORMATION: R group has a pyridine ring.  
US-08-446-206B-22

Query Match

Best Local Similarity 21.4%; Score 3; DB 1; Length 8;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 GXGV 8

DB 4 GXGV 7

RESULT 13

US-08-446-206B-28

Sequence 28, Application US/08446206B

Patent No. 5750647

GENERAL INFORMATION:

APPLICANT: Eyre, David R

APPLICANT: Clemens, J Daniel

APPLICANT: Ochs, Vincent W

TITLE OF INVENTION: Synthetic Peptide Analogs of NTX

NUMBER OF SEQUENCES: 31

CORRESPONDENCE ADDRESS:

ADDRESSEE: Christensen, O'Connor, Johnson & Kindness PLLC

STREET: 1420 Fifth Avenue

CITY: Seattle

STATE: Washington

COUNTRY: U.S.A.

ZIP: WA 98101

COMPUTER READABLE FORM:

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/446,206B

FILING DATE: 19-MAY-1995

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Shelton, Dennis K

REGISTRATION NUMBER: 26,997

REFERENCE/DOCKET NUMBER: WROS18360

TELECOMMUNICATION INFORMATION:

TELEPHONE: 206 682 8100

TELEFAX: 206 224 0779

TELEX:

INFORMATION FOR SEQ ID NO: 28:

SEQUENCE CHARACTERISTICS:

LENGTH: 8 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: peptide

FRAGMENT TYPE: N-terminal

FEATURE:

LOCATION: 1

OTHER INFORMATION: Xaa is Gln is glutamine or

OTHER INFORMATION: pyrrolidone carboxylic acid.  
 FEATURE: 5  
 LOCATION: 5  
 OTHER INFORMATION: Xaa is an alpha-amino acid bearing a  
 OTHER INFORMATION: thiol group.  
 US-08-446-206B-28

Query Match 21.4%; Score 3; DB 1; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+05;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 GXGV 8  
 DB 4 GXGV 7

RESULT 14  
 US-08-446-206B-29  
 ; Sequence 29, Application US/08446206B  
 ; Patent No. 5750647  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Eyre, David R  
 ; APPLICANT: Clemens, J Daniel  
 ; APPLICANT: Ochs, Vincent W  
 ; TITLE OF INVENTION: Synthetic Peptide Analogs of NTx  
 ; NUMBER OF SEQUENCES: 31  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Christensen, O'Connor, Johnson & Kindness PLLC  
 ; STREET: 1420 Fifth Avenue  
 ; CITY: Seattle  
 ; STATE: Washington  
 ; COUNTRY: U.S.A.  
 ; ZIP: WA 98101  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/446,206B  
 ; FILING DATE: 19-MAY-1995  
 ; CLASSIFICATION: 530  
 ; PRIOR APPLICATION NUMBER:  
 ; FILING DATE:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Shelton, Dennis K  
 ; REGISTRATION NUMBER: 26,997  
 ; REFERENCE/DOCKET NUMBER: WRO518360  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 206 682 8100  
 ; TELEFAX: 206 224 0779  
 ; TELEX:  
 ; INFORMATION FOR SEQ ID NO: 29:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 8 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS:  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: peptide  
 ; FRAGMENT TYPE: N-terminal  
 ; FEATURE:  
 ; LOCATION: 1  
 ; OTHER INFORMATION: Xaa is glutamine or pyrrolidone  
 ; OTHER INFORMATION: carboxylic acid.  
 ; FEATURE:  
 ; LOCATION: 5  
 ; OTHER INFORMATION: Xaa is homocysteine.  
 ; OTHER INFORMATION:  
 ; US-08-446-206B-29

Query Match 21.4%; Score 3; DB 1; Length 8;

Best Local Similarity 100.0%; Pred. No. 1.7e+05;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 GXGV 8  
 DB 4 GXGV 7

RESULT 15  
 US-08-446-206B-30  
 ; Sequence 30, Application US/08446206B  
 ; Patent No. 5750647  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Eyre, David R  
 ; APPLICANT: Clemens, J Daniel  
 ; APPLICANT: Ochs, Vincent W  
 ; TITLE OF INVENTION: Synthetic Peptide Analogs of NTx  
 ; NUMBER OF SEQUENCES: 31  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Christensen, O'Connor, Johnson & Kindness PLLC  
 ; STREET: 1420 Fifth Avenue  
 ; CITY: Seattle  
 ; STATE: Washington  
 ; COUNTRY: U.S.A.  
 ; ZIP: WA 98101  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/446,206B  
 ; FILING DATE: 19-MAY-1995  
 ; CLASSIFICATION: 530  
 ; PRIOR APPLICATION NUMBER:  
 ; FILING DATE:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Shelton, Dennis K  
 ; REGISTRATION NUMBER: 26,997  
 ; REFERENCE/DOCKET NUMBER: WRO518360  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 206 682 8100  
 ; TELEFAX: 206 224 0779  
 ; TELEX:  
 ; INFORMATION FOR SEQ ID NO: 30:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 8 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS:  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: peptide  
 ; FRAGMENT TYPE: N-terminal  
 ; FEATURE:  
 ; LOCATION: 1  
 ; OTHER INFORMATION: Xaa is glutamine or pyrrolidone  
 ; OTHER INFORMATION: carboxylic acid.  
 ; FEATURE:  
 ; LOCATION: 5  
 ; OTHER INFORMATION: Xaa is an alpha-amino acid bearing a  
 ; OTHER INFORMATION: side chain amino group or a side chain thiol  
 ; OTHER INFORMATION: group.  
 ; US-08-446-206B-30

Query Match 21.4%; Score 3; DB 1; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+05;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 GXGV 8  
 DB 4 GXGV 7

Search completed: October 4, 2002, 10:30:09  
Job time: 226 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 4, 2002, 10:30:40 ; Search time 24.71 Seconds  
(without alignments)  
54.442 Million cell updates/sec

Title: US-09-822-698A-28  
Perfect score: 14  
Sequence: 1 XXHTGQVWXPXXX 14

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 283138 seqs, 96089334 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : PIR71:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3	21.4	8	2 A54823	olfactory receptor
2	3	21.4	13	2 PS0250	ribulose-bisphosph
3	3	21.4	14	2 S57638	T cell receptor V-
4	3	21.4	14	2 D45474	thrombospondin 2 -
5	3	21.4	15	2 A60929	dichloromethane de
6	3	21.4	15	2 B60929	dichloromethane de
7	3	21.4	16	2 D45193	zinc finger protei
8	3	21.4	20	2 A60801	acrosome stabilizi
9	3	21.4	20	2 PQ0537	arylhydroxamic aci
10	3	21.4	25	2 JA0171	dnak-type molecula
11	3	21.4	28	2 S47614	zinc finger protei
12	3	21.4	28	2 I39288	ZF3 domain - human
13	3	21.4	29	2 F42075	finger protein (cl
14	3	21.4	30	2 A31187	neurotoxin II-22.5
15	3	21.4	31	2 S18356	chymotrypsin (EC 3
16	3	21.4	32	2 S30758	genome polyprotein
17	3	21.4	32	2 G42075	finger protein (cl
18	3	21.4	36	2 F45193	zinc finger protei
19	3	21.4	37	2 AH2787	hypothetical prote
20	3	21.4	39	2 S71913	hemoglobin, extrac
21	3	21.4	39	2 I65269	collagen alpha 1(I
22	3	21.4	39	2 C40984	finger protein 18
23	3	21.4	39	2 D40984	finger protein zfa
24	3	21.4	39	2 B40984	finger protein zfe
25	3	21.4	40	2 A46041	zinc-proteinase -
26	3	21.4	43	2 B42825	Kruppel-type zinc
27	3	21.4	43	2 E64740	hypothetical prote
28	3	21.4	44	2 I37956	zinc finger protei
29	3	21.4	45	2 A84143	hypothetical prote

```

30      3      21.4      46      2      A61078      glucose-6-phosphat
31      3      21.4      46      2      E90676      hypothetical prote
32      3      21.4      46      2      H85526      hypothetical prote
33      3      21.4      48      2      H90799      hypothetical prote
34      3      21.4      49      2      S72972      hypothetical prote
35      3      21.4      50      2      I37968      zinc finger protei
36      3      21.4      51      2      PN0479      cell division cycl
37      3      21.4      52      2      S45374      translation elonga
38      3      21.4      52      2      B90673      probable invertase
39      3      21.4      52      2      S72985      glutamate decarbox
40      3      21.4      52      2      F43284      finger protein ZNF
41      3      21.4      52      2      I37971      zinc finger protei
42      3      21.4      52      2      C38203      clone GT4-1 - mous
43      3      21.4      52      2      AB0004      probable membrane
44      3      21.4      52      2      F85523      probable invertase
45      3      21.4      53      2      S13105      cytochrome-c oxida

```

#### ALIGNMENTS

```

RESULT 1
A54823
olfactory receptor I7 - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 28-Apr-1995 #sequence_revision 28-Apr-1995 #text_change 17-Mar-1999
C:Accession: A54823
R:Chess, A.; Simon, I.; Cedar, H.; Axel, R.
Cell 78, 823-834, 1994
A:Title: Allelic inactivation regulates olfactory receptor gene expression.
A:Reference number: A54823; MUID:94373818
A:Accession: A54823
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-8 <CHE>

```

```

Query Match      21.4%; Score 3; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      3 HTG 5
      |||
DB      6 HTG 8

```

```

RESULT 2
PS0250
ribulose-bisphosphate carboxylase activase II - rice (strain Nihonbare) (fragment)
C:Species: Oryza sativa (rice)
C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 02-Jul-1998
C:Accession: PS0250
R:Tsugita, A.
submitted to JIPID, April 1993
A:Reference number: PS0206
A:Accession: PS0250
A:Molecule type: protein
A:Residues: 1-13 <TSU>
A:Experimental source: leaf, chlorophyll

```

```

Query Match      21.4%; Score 3; DB 2; Length 13;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      4 TCGX 7
      ||||
DB      5 TCGX 8

```

```

RESULT 3
S57638
T cell receptor V-J junctional alpha chain region - human (fragment)

```

C:Species: Homo sapiens (man)  
C:Date: 19-Oct-1995 #sequence\_revision 17-Nov-1995 #text\_change 05-Nov-1999  
C:Accession: S57638  
R:Burrows, S.R.; Sillins, S.L.; Moss, D.J.; Khanna, R.; Misko, I.S.; Arguet, V.P.  
submitted to the EMBL Data Library, June 1995  
A:Description: T cell receptor repertoire for a viral epitope in humans is diversified b  
A:Reference number: S57494  
A:Accession: S57638  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-14 <BUR>  
A:Cross-references: EMBL:Z49964; NID:g886676; PIDN:CAA90238.1; PID:g886677  
C:Keywords: T-cell receptor

Query Match 21.4%; Score 3; DB 2; Length 14;  
Best Local Similarity 100.0%; Pred. No. 2.8e+02;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 HTG 5

Db 6 HTG 8

RESULT 4

D45474

thrombospondin 2 - bovine (fragment)  
N:Alternate names: corticotropin-induced secreted protein (CISP); thrombospondin homolog  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 21-Sep-1993 #sequence\_revision 18-Nov-1994 #text\_change 28-Feb-1997  
C:Accession: D45474  
R:Pellerin, S.; Lafeuillade, B.; Scherrer, N.; Gagnon, J.; Shi, D.L.; Chambaz, E.M.; Fe  
J. Biol. Chem. 268, 4304-4310, 1993  
A:Title: Corticotropin-induced secreted protein, an ACTH-induced protein secreted by ad  
A:Reference number: A45474; MUID:93179438  
A:Accession: D45474  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-14 <PEL>  
A:Experimental source: adrenocortical cells  
A:Note: sequence extracted from NCBI backbone (NCBIP:125846)

Query Match 21.4%; Score 3; DB 2; Length 14;  
Best Local Similarity 100.0%; Pred. No. 2.8e+02;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 HTG 5

Db 5 HTG 7

RESULT 5

A60929

dichloromethane dehalogenase (EC 4.5.1.3) - Hyphomicrobium sp. (ATCC 43129) (fragment)  
C:Species: Hyphomicrobium sp.  
C:Date: 03-Feb-1994 #sequence\_revision 03-Feb-1994 #text\_change 07-Oct-1994  
C:Accession: A60929  
R:Kohler-Staub, D.; Hartmans, S.; Gaelli, R.; Suter, F.; Leisinger, T.  
J. Gen. Microbiol. 132, 2837-2843, 1986  
A:Title: Evidence for identical dichloromethane dehalogenases in different methylotroph  
A:Reference number: A60929  
A:Accession: A60929  
A:Molecule type: protein  
A:Residues: 1-15 <KOH>  
C:Keywords: carbon-halide lyase

Query Match 21.4%; Score 3; DB 2; Length 15;  
Best Local Similarity 100.0%; Pred. No. 3e+02;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 HTG 5

Db 9 HTG 11

RESULT 6

B60929

dichloromethane dehalogenase (EC 4.5.1.3) - Pseudomonas sp. (fragment)  
C:Species: Pseudomonas sp.  
C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 07-Oct-1994  
C:Accession: B60929  
R:Kohler-Staub, D.; Hartmans, S.; Gaelli, R.; Suter, F.; Leisinger, T.  
J. Gen. Microbiol. 132, 2837-2843, 1986  
A:Title: Evidence for identical dichloromethane dehalogenases in different methylotro  
A:Reference number: A60929  
A:Contents: DSM 1565  
A:Accession: B60929  
A:Molecule type: protein  
A:Residues: 1-15 <KOH>  
C:Keywords: carbon-halide lyase

Query Match 21.4%; Score 3; DB 2; Length 15;  
Best Local Similarity 100.0%; Pred. No. 3e+02;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 HTG 5

Db 9 HTG 11

RESULT 7

D45193

zinc finger protein ZNF60 - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 01-Jan-1993 #sequence\_revision 18-Nov-1994 #text\_change 05-Nov-1999  
C:Accession: D45193  
R:Lichter, P.; Bray, P.; Ried, T.; David, I.B.; Ward, D.C.  
Genomics 13, 999-1007, 1992  
A:Title: Clustering of C2-H2 zinc finger motif sequences within telomeric and fragile  
A:Reference number: A43284; MUID:92372070  
A:Accession: D45193  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-16 <LIC>  
A:Cross-references: GB:M88369; NID:g340475; PIDN:AAA61327.1; PID:g340476  
A:Note: sequence extracted from NCBI backbone (NCBIN:111662, NCBIP:111664)

Query Match 21.4%; Score 3; DB 2; Length 16;  
Best Local Similarity 100.0%; Pred. No. 3.1e+02;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 HTG 5

Db 2 HTG 4

RESULT 8

A60801

acrosome stabilizing factor large chain - rabbit (fragment)  
C:Species: Oryctolagus cuniculus (domestic rabbit)  
C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 07-May-1999  
C:Accession: A60801  
R:Wilson, W.L.; Oliphant, G.  
Biol. Reprod. 37, 159-169, 1987  
A:Title: Isolation and biochemical characterization of the subunits of the rabbit spe  
A:Reference number: A60801; MUID:88000873  
A:Accession: A60801  
A:Molecule type: protein  
A:Residues: 1-20 <WIL>  
C:Comment: spermatozoa must undergo capacitation and the acrosome reaction to become  
C:Keywords: glycoprotein; semen



Query Match 21.4%; Score 3; DB 2; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 3.8e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 TGXG 7  
 ||||  
 Db 4 TGXG 7

RESULT 9  
 PQ0537  
 arylhydroxamic acid N,O-acetyltransferase (EC 2.3.1.-) - golden hamster (fragment)  
 C:Species: Mesocricetus auratus (golden hamster)  
 C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 07-May-1999  
 C:Accession: PQ0537  
 A:Source, Pharm. Bull. 40, 2857-2859, 1992  
 A:Title: Purification and characterization of hamster hepatic microsomal N,O-acetyltransferase  
 A:Reference number: PQ0537; MUID:93099627  
 A:Accession: PQ0537  
 A:Molecule type: protein  
 A:Residues: 1-20 <SON>  
 A:Experimental source: liver  
 C:Keywords: acyltransferase

Query Match 21.4%; Score 3; DB 2; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 3.8e+02;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 HTG 5  
 ||||  
 Db 10 HTG 12

RESULT 10  
 JA0171  
 dnaK-type molecular chaperone 70-3 - Arabidopsis thaliana (fragment)  
 N:Alternate names: heat shock protein 70-3  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 12-Mar-1993 #sequence\_revision 12-Mar-1993 #text\_change 13-Mar-1998  
 C:Accession: JA0171  
 R:Wu, C.H.; Caspar, T.; Browne, J.; Lindquist, S.; Somerville, C.  
 Plant Physiol. 86, 731-740, 1988  
 A:Title: Characterization of an HSP70 cognate gene family in Arabidopsis.  
 A:Reference number: JA0169  
 A:Accession: JA0171  
 A:Molecule type: DNA  
 A:Residues: 1-25 <WUC>  
 A:Cross-references: GB:M23107  
 A:Experimental source: leaf  
 C:Genetics:  
 A:Introns: 25/3  
 C:Function:  
 A:Description: involved in protein folding and assembling/disassembling of protein complex  
 C:Superfamily: heat shock protein 70  
 C:Keywords: Arp; heat shock; molecular chaperone; stress-induced protein

Query Match 21.4%; Score 3; DB 2; Length 25;  
 Best Local Similarity 100.0%; Pred. No. 4.6e+02;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 GWV 9  
 ||||  
 Db 22 GWV 24

RESULT 11  
 S47614  
 zinc finger protein 11 - mouse (fragment)  
 C:Species: Mus musculus (house mouse)

C:Date: 07-May-1995 #sequence\_revision 24-Oct-1998 #text\_change 03-Nov-2000  
 C:Accession: S47614  
 R:Stone, B.; Wharton, W.  
 Nucleic Acids Res. 22, 2612-2618, 1994  
 A:Title: Targeted RNA fingerprinting: the cloning of differentially-expressed cDNA f.  
 A:Reference number: S47614; MUID:94316505  
 A:Accession: S47614  
 A:Molecule type: mRNA  
 A:Residues: 1-28 <STO>  
 C:Superfamily: transcription factor Krueppel  
 C:Keywords: zinc finger

Query Match 21.4%; Score 3; DB 2; Length 28;  
 Best Local Similarity 100.0%; Pred. No. 5e+02;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 HTG 5  
 ||||  
 Db 1 HTG 3

RESULT 12  
 I39288  
 ZF3 domain - human (fragment)  
 C:Species: Homo sapiens (man)  
 C:Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 05-Nov-1999  
 C:Accession: I39288  
 R:Osawa, O.; Eccles, M.R.; Yun, K.; Mueller, R.F.; Holdaway, M.D.; Reeve, A.E.  
 Hum. Mol. Genet. 2, 203-204, 1993  
 A:Title: A novel insertional mutation at the third zinc finger coding region of the W  
 A:Reference number: I39288; MUID:93271969  
 A:Accession: I39288  
 A:Status: preliminary; translated from GB/EMBL/DBBJ  
 A:Molecule type: DNA  
 A:Residues: 1-28 <RES>  
 A:Cross-references: EMBL:X72314; NID:g312849; PIDN:CAA51057.1; PID:g312850  
 C:Genetics:  
 A:Gene: GDB:WT1  
 A:Cross-references: GDB:I20496; OMIM:194070  
 A:Map position: 11p13-11p13

Query Match 21.4%; Score 3; DB 2; Length 28;  
 Best Local Similarity 100.0%; Pred. No. 5e+02;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 HTG 5  
 ||||  
 Db 22 HTG 24

RESULT 13  
 F42075  
 finger protein (clone Znpf30) - human (fragment)  
 C:Species: Homo sapiens (man)  
 C:Date: 28-Apr-1993 #sequence\_revision 28-Apr-1993 #text\_change 01-Dec-2000  
 C:Accession: F42075  
 R:Hoovers, J.M.N.; Mannens, M.; John, R.; Bliet, J.; van Heyningen, V.; Porteous, D.J.  
 Genomics 12, 254-263, 1992  
 A:Title: High-resolution localization of 69 potential human zinc finger protein genes  
 A:Reference number: A42075; MUID:92155714  
 A:Accession: F42075  
 A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual  
 A:Molecule type: DNA  
 A:Residues: 1-29 <HOO>  
 C:Superfamily: zinc finger protein ZFP-36; LIM metal-binding repeat homology

Query Match 21.4%; Score 3; DB 2; Length 29;  
 Best Local Similarity 100.0%; Pred. No. 5.2e+02;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Fri Oct 4 11:00:23 2002

OY 3 HTG 5  
 |||  
 Db 15 HTG 17

## RESULT 14

A31187  
 neurotoxin II.22.5 - scorpion (Centruroides limpidus) (fragment)  
 C:Species: Centruroides limpidus tecomanus  
 C:Date: 28-Feb-1990 #sequence\_revision 28-Feb-1990 #text\_change 07-Feb-1997  
 C:Accession: A31187  
 R:Ramirez, A.N.; Gurrola, G.B.; Martin, B.M.; Possani, L.D.  
 Toxicon 26, 773-783, 1988  
 A:Title: Isolation of several toxins from the venom of the scorpion Centruroides limpidus  
 A:Reference number: A31187; MUID:89073188  
 A:Accession: A31187  
 A:Molecule type: protein  
 A:Residues: 1-30 <RAM>  
 C:Superfamily: scorpion neurotoxin  
 C:Keywords: neurotoxin

Query Match 21.4%; Score 3; DB 2; Length 30;  
 Best Local Similarity 100.0%; Pred. No. 5.3e+02;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 HTG 5  
 |||  
 Db 9 HTG 11

## RESULT 15

S18356  
 chymotrypsin (EC 3.4.21.1) Pm2 - penaeid shrimp (Penaeus monodon) (fragment)  
 C:Species: Penaeus monodon  
 C:Date: 22-Nov-1993 #sequence\_revision 01-Mar-1996 #text\_change 01-Nov-1996  
 C:Accession: S18356  
 R:Tsai, I.H.; Lu, P.J.; Chuang, J.L.  
 Biochim. Biophys. Acta 1080, 59-67, 1991  
 A:Title: The midgut chymotrypsins of shrimps (Penaeus monodon, Penaeus japonicus and Penaeus chinensis)  
 A:Reference number: S18356; MUID:92031652  
 A:Accession: S18356  
 A:Molecule type: protein  
 A:Residues: 1-31 <TSA>  
 C:Superfamily: trypsin; trypsin homology  
 C:Keywords: hydrolase; protein digestion; serine proteinase

Query Match 21.4%; Score 3; DB 2; Length 31;  
 Best Local Similarity 100.0%; Pred. No. 5.5e+02;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 GWV 9  
 |||  
 Db 10 GWV 12

Search completed: October 4, 2002, 10:30:41  
 Job time: 183 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 4, 2002, 10:36:36 ; Search time 15.75 Seconds  
(without alignments)  
34.417 Million cell updates/sec

Title: US-09-822-698A-28  
Perfect score: 14  
Sequence: 1 XXHTGXGVWXPXXX 14

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 105224 seqs, 38719550 residues

Word size : 0

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	3	21.4	10	1 LABA_JATMU	P13270 jatropha mu
2	3	21.4	19	1 ETFA_CLOPA	P81342 clostridium
3	3	21.4	30	1 SCX2_CENLI	P18927 centruridae
4	3	21.4	31	1 CTRP_PENMO	P35002 penaeus mon
5	3	21.4	46	1 ESCA_CRYAP	P31509 cryptops ap
6	3	21.4	46	1 ESCA_LITFO	Q02025 lithobius f
7	3	21.4	46	1 ESCA_ORVLA	Q01798 oryzias lat
8	3	21.4	46	1 ESCA_PHOPA	Q02026 pholcus pha
9	3	21.4	46	1 ESCA_PSYCI	Q02027 psychoda ci
10	3	21.4	46	1 ESCA_TRICA	Q01800 tribolium c
11	3	21.4	46	1 KRUP_CUPSA	Q02415 cupiennius
12	3	21.4	46	1 KRUP_LITFO	Q01872 lithobius f
13	3	21.4	46	1 KRUP_PHOPA	Q02034 pholcus pha
14	3	21.4	50	1 LACA_LACAM	P80696 lactobacill
15	3	21.4	50	1 ZN30_HUMAN	P17039 homo sapien
16	3	21.4	52	1 COXE_DICDI	P20609 dictyosteli
17	3	21.4	52	1 ZN32_HUMAN	P17041 homo sapien
18	3	21.4	53	1 MT_SYNBP	P08002 synchococc
19	3	21.4	54	1 YJUX_ENTAE	P39430 enterobacte
20	3	21.4	55	1 MT_SYNBP7	P30331 synchococc
21	3	21.4	55	1 R29A_YEAST	P41057 saccharomyc
22	3	21.4	55	1 R29B_YEAST	P41058 saccharomyc
23	3	21.4	55	1 RS29_SCHPO	O74329 schizosacch
24	3	21.4	56	1 KRXA_MOUSE	P17051 mus musculu
25	3	21.4	56	1 ZK23_HUMAN	P17034 homo sapien
26	3	21.4	56	1 ZN12_HUMAN	P17014 homo sapien
27	3	21.4	56	1 ZN13_HUMAN	P17016 homo sapien
28	3	21.4	56	1 ZN14_HUMAN	P17017 homo sapien
29	3	21.4	56	1 ZN15_HUMAN	P17019 homo sapien
30	3	21.4	56	1 ZN16_HUMAN	P17020 homo sapien
31	3	21.4	56	1 ZN17_HUMAN	P17021 homo sapien
32	3	21.4	56	1 ZN18_HUMAN	P17022 homo sapien
33	3	21.4	56	1 ZN21_HUMAN	P17025 homo sapien

34	3	21.4	56	1 ZN25_HUMAN	P17030 homo sapien
35	3	21.4	56	1 ZN26_HUMAN	P17031 homo sapien
36	3	21.4	56	1 ZN27_HUMAN	P17033 homo sapien
37	3	21.4	56	1 ZN29_HUMAN	P17037 homo sapien
38	3	21.4	56	1 ZN31_HUMAN	P17040 homo sapien
39	3	21.4	57	1 GRN1_CYPCA	P81013 cyprinus ca
40	3	21.4	57	1 GRN2_CYPCA	P81014 cyprinus ca
41	3	21.4	57	1 GRN3_CYPCA	P81015 cyprinus ca
42	3	21.4	57	1 YCU5_CAEEL	Q22702 caenorhabdi
43	3	21.4	59	1 EGR1_SERCA	O73694 serinus can
44	3	21.4	59	1 FOAH_SALTY	P56505 salmonella
45	3	21.4	60	1 TAP_ORNMO	P17726 ornithodoros

#### ALIGNMENTS

RESULT 1  
LABA\_JATMU STANDARD; PRT; 10 AA.  
ID LABA\_JATMU  
AC P13270;  
DT 01-JAN-1990 (Rel. 13, Created)  
DT 01-JAN-1990 (Rel. 13, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE Labaditin.  
OS Jatropha multifida (Physic nut).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids I; Malpighiales; Euphorbiaceae; Jatropha.  
OX NCBI\_TaxID=3996;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Latex;  
RA Kosasi S., van der Sluis W.G., Boelens R., T'Harb L.A., Labadie R.P.;  
RT "Labaditin, a novel cyclic decapeptide from the latex of Jatropha  
multifida L. (Euphorbiaceae). Isolation and sequence determination  
by means of two-dimensional NMR.";  
RL FEBS Lett. 256:91-95(1989).  
CC -!- FUNCTION: LABADITIN IS AN ACTIVE PEPTIDE WHICH INHIBITS THE  
CLASSICAL PATHWAY OF COMPLEMENT ACTIVATION IN VITRO. ACTIVITY  
SEEMS TO BE BASED ON AN INTERACTION WITH C1.  
CC -!- DISEASE: LATEX OF THIS PLANT IS USED IN FOLKLORE MEDICINE FOR  
TREATMENT OF INFECTED WOUNDS, SKIN INFECTIONS AND SCABIES.  
CC -!- CAUTION: THIS IS A CYCLIC PEPTIDE.  
KW Latex.  
SQ SEQUENCE 10 AA; 1089 MW; D98AAD6362D1B362 CRC64;

Query Match 21.48; Score 3; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 89;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 GWV 9  
|||  
Db 2 GWV 4

RESULT 2  
ETFA\_CLOPA STANDARD; PRT; 19 AA.  
ID ETFA\_CLOPA  
AC P81342;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE Electron transfer flavoprotein alpha-subunit (Alpha-ETF) (Electron  
transfer flavoprotein large subunit) (ETF-L) (CP 14) (Fragment).  
ETFA.  
OS Clostridium pasteurianum.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;  
OC Clostridium.  
OX NCBI\_TaxID=1501;  
RN [1]  
RP SEQUENCE.

## RESULT 4

RESULT 5

ESCA_CRYAP	STANDARD;	PRT;	46 AA.
ID			
AC	P31509;		
DT	01-JUL-1993 (Rel. 26, Created)		
DT	01-JUL-1993 (Rel. 26, Last sequence update)		
DT	16-OCT-2001 (Rel. 40, Last annotation update)		
DE	Escargot/snail protein homolog (Fragment).		
OS	Cryptops apomaterials.		
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Myriapoda; Chilopoda;		
QC	Pleurostigmophora; Scolopendromorpha; Cryptopsidae; Cryptops.		
OX	NCBI_TAXID=7545;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=93066327; PubMed=1438276;		
RA	Sommer R.J., Retzlaff M., Sander K., Tautz D.;		
RT	"Evolutionary conservation pattern of zinc-finger domains of		
RL	Drosophila segmentation genes.";		
RT	Proc. Natl. Acad. Sci. U.S.A. 89:10782-10786(1992).		
CC	-I- SUBCELLULAR LOCATION: Nuclear (Probable).		
CC	-I- SIMILARITY: BELONGS TO THE SNAIL FAMILY OF ZINC FINGER PROTEINS.		
CC			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		

CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch))

DR EMBL; L01593; AAA28267.1; -;  
DR HSP; P08046; IAILL.  
DR InterPro; IPR000822; Znf-C2H2.  
DR Pfam; PF00096; zf-C2H2; 1.  
DR SMART; SM00355; Znf\_C2H2; 1.  
DR PROSITE; PS00028; ZINC\_FINGER\_C2H2\_1; 1.  
DR PROSITE; PS00157; ZINC\_FINGER\_C2H2\_2; 1.  
KW Zinc-finger; Metal-binding; DNA-binding; Repeat; Nuclear protein.  
FT DOMAIN 1 1  
FT ZN\_FING <1 >46 ZINC-FINGERS.  
FT ZN\_FING <1 4 C2H2-TYPE.  
FT ZN\_FING 8 30 C2H2-TYPE.  
FT ZN\_FING 36 >46 C2H2-TYPE.  
FT NON\_TER 46 46  
SQ SEQUENCE 46 AA; 5360 MW; 22F14B5C16EADB10 CRC64;

Query Match 21.4%; Score 3; DB 1; Length 46;  
Best Local Similarity 100.0%; Pred. No. 3.2e+02;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 HTG 5  
|||  
Db 30 HTG 32

RESULT 6  
ESCA\_LITFO STANDARD; PRT; 46 AA.  
AC Q02025;  
DT 01-JUL-1993 (Rel. 26, Created)  
DT 01-JUL-1993 (Rel. 26, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Escargot/snail protein homolog (Fragment).  
OS Lithobius forficatus.  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Myriapoda; Chilopoda;  
OC Pleurostigmophora; Lithobiomorpha; Lithobiidae; Lithobius.  
OX NCBI\_TaxID=7552;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93066327; PubMed=1438276;  
RA Sommer R.J., Retzlaff M., Goerlich K., Sander K., Tautz D.;  
RT "Evolutionary conservation pattern of zinc-finger domains of  
Drosophila segmentation genes."  
RL Proc. Natl. Acad. Sci. U.S.A. 89:10782-10786(1992).  
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).  
CC -1- SIMILARITY: BELONGS TO THE SNAIL FAMILY OF ZINC FINGER PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch))

DR EMBL; L01599; AAA29281.1; -;  
DR HSP; P07248; 2ADR  
DR InterPro; IPR000822; Znf-C2H2.  
DR Pfam; PF00096; zf-C2H2; 1.  
DR SMART; SM00355; Znf\_C2H2; 1.  
DR PROSITE; PS00028; ZINC\_FINGER\_C2H2\_1; PARTIAL.  
DR PROSITE; PS00157; ZINC\_FINGER\_C2H2\_2; 1.  
KW Zinc-finger; Metal-binding; DNA-binding; Repeat; Nuclear protein.  
FT DOMAIN 1 1  
FT ZN\_FING <1 >46 ZINC-FINGERS.  
FT ZN\_FING <1 4 C2H2-TYPE.  
FT ZN\_FING 9 30 C2H2-TYPE.  
FT ZN\_FING 36 >46 C2H2-TYPE.

FT NON\_TER 46  
SQ SEQUENCE 46 AA; 5102 MW; 1EF6A856A394C640 CP:54;

Query Match 21.4%; Score 3; DB 1; Length 46;  
Best Local Similarity 100.0%; Pred. No. 3.2e+02;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 HTG 5  
|||  
Db 30 HTG 32

RESULT 7  
ESCA\_ORYLA STANDARD; PRT; 46 AA.  
ID ESCA\_ORYLA  
AC Q01798;  
DT 01-JUL-1993 (Rel. 26, Created)  
DT 01-JUL-1993 (Rel. 26, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Escargot/snail protein homolog (Fragment).  
OS Oryzias latipes (Medaka fish).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;  
OC Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.  
OX NCBI\_TaxID=8090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93066327; PubMed=1438276;  
RA Sommer R.J., Retzlaff M., Goerlich K., Sander K., Tautz D.;  
RT "Evolutionary conservation pattern of zinc-finger domains of  
Drosophila segmentation genes."  
RL Proc. Natl. Acad. Sci. U.S.A. 89:10782-10786(1992).  
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).  
CC -1- SIMILARITY: BELONGS TO THE SNAIL FAMILY OF ZINC FINGER PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch))

DR EMBL; L01603; AAA49441.1; -;  
DR HSP; P03001; ITF6  
DR InterPro; IPR000822; Znf-C2H2.  
DR Pfam; PF00096; zf-C2H2; 1.  
DR SMART; SM00355; Znf\_C2H2; 1.  
DR PROSITE; PS00028; ZINC\_FINGER\_C2H2\_1; 1.  
DR PROSITE; PS00157; ZINC\_FINGER\_C2H2\_2; 1.  
KW Zinc-finger; Metal-binding; DNA-binding; Repeat; Nuclear protein.  
FT NON\_TER 1 1  
FT DOMAIN <1 >46 ZINC-FINGERS.  
FT ZN\_FING <1 4 C2H2-TYPE.  
FT ZN\_FING 8 30 C2H2-TYPE.  
FT ZN\_FING 36 >46 C2H2-TYPE.  
FT NON\_TER 46 46  
SQ SEQUENCE 46 AA; 5218 MW; 22EDD03517345B10 CRC64;

Query Match 21.4%; Score 3; DB 1; Length 46;  
Best Local Similarity 100.0%; Pred. No. 3.2e+02;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 HTG 5  
|||  
Db 30 HTG 32

RESULT 8  
ESCA\_PHOPA

```

ID  ESCA_PHOPA  STANDARD;  PRT;  46 AA.
AC  Q02026;
DT  01-JUL-1993 (Rel. 26, Created)
DT  01-JUL-1993 (Rel. 26, Last sequence update)
DE  16-OCT-2001 (Rel. 40, Last annotation update)
DE  Escargot/snail protein homolog (Fragment).
OS  Pholcus phalangioidea.
OC  Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;
OC  Araneomorphae; Haplogynae; Pholcidae; Pholcus.
OX  NCBI_TaxID=6932;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=93066327; PubMed=1438276;
RA  Sommer R.J., Retzlaff M., Goerlich K., Sander K., Tautz D.;
RT  "Evolutionary conservation pattern of zinc-finger domains of
RT  Drosophila segmentation genes.";
RL  Proc. Natl. Acad. Sci. U.S.A. 89:10782-10786(1992).
CC  -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC  -!- SIMILARITY: BELONGS TO THE SNAIL FAMILY OF ZINC FINGER PROTEINS.
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; L01610; AAA29800.1; -
DR  HSSP; P08046; 1A1L.
DR  InterPro; IPR000822; Znf-C2H2.
DR  Pfam; PF00096; zf-C2H2; 1.
DR  SMART; SM00355; Znf_C2H2; 1.
DR  PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
DR  PROSITE; PS0157; ZINC_FINGER_C2H2_2; 1.
KW  Zinc-finger; Metal-binding; DNA-binding; Repeat; Nuclear protein.
FT  DOMAIN 1 <1 >46 ZINC-FINGERS.
FT  ZN_FING <1 >4 C2H2-TYPE.
FT  ZN_FING 8 30 C2H2-TYPE.
FT  ZN_FING 36 >46 C2H2-TYPE.
FT  NON_TER 46 46
SQ  SEQUENCE 46 AA; 5330 MW; 22E3FCAC16EADB10 CRC64;

Query Match 21.4%; Score 3; DB 1; Length 46;
Best Local Similarity 100.0%; Pred. No. 3.2e-02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 HTG 5
DB 30 HTG 32

RESULT 10
ESCA_TRICA  STANDARD;  PRT;  46 AA.
ID  ESCA_TRICA  STANDARD;  PRT;  46 AA.
AC  Q01800;
DT  01-JUL-1993 (Rel. 26, Created)
DT  01-JUL-1993 (Rel. 26, Last sequence update)
DT  16-OCT-2001 (Rel. 40, Last annotation update)
DE  Escargot/snail protein homolog (Fragment).
OS  Tribolium castaneum (Red flour beetle).
OC  Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC  Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
OC  Cucujiformia; Tenebrionidae; Tribolium.
OX  NCBI_TaxID=7070;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=93066327; PubMed=1438276;
RA  Sommer R.J., Retzlaff M., Goerlich K., Sander K., Tautz D.;
RT  "Evolutionary conservation pattern of zinc-finger domains of
RT  Drosophila segmentation genes.";
RL  Proc. Natl. Acad. Sci. U.S.A. 89:10782-10786(1992).
CC  -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC  -!- SIMILARITY: BELONGS TO THE SNAIL FAMILY OF ZINC FINGER PROTEINS.
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; L01617; AAA30097.1; -
DR  HSSP; P08047; 1SP2.
DR  InterPro; IPR000822; Znf-C2H2.
DR  Pfam; PF00096; zf-C2H2; 1.

```

DR SMART: SM00355; ZnF\_C2H2; 1.  
 DR PROSITE: PS00028; ZINC\_FINGER\_C2H2\_1; 1.  
 DR PROSITE: PS50157; ZINC\_FINGER\_C2H2\_2; 1.  
 KW Zinc-finger; Metal-binding; DNA-binding; Repeat; Nuclear protein.  
 FT NON\_TER 1 1  
 FT DOMAIN <1 >46 ZINC-FINGERS.  
 FT ZN\_FING <1 4 C2H2-TYPE.  
 FT ZN\_FING 8 30 C2H2-TYPE.  
 FT ZN\_FING 37 >46 C2H2-TYPE.  
 FT NON\_TER 46 46  
 SQ SEQUENCE 46 AA; 5292 MW; 7D193BAD416E81A7 CRC64;

Query Match 21.4%; Score 3; DB 1; Length 46;  
 Best Local Similarity 100.0%; Pred. No. 3.2e+02;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 HTG 5  
 Db 30 HTG 32

RESULT 11  
 KRUP\_CUPSA  
 ID KRUP\_CUPSA STANDARD; PRT; 46 AA.  
 AC Q02415;  
 DT 01-JUL-1993 (Rel. 26, Created)  
 DT 01-JUL-1993 (Rel. 26, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Krueppel protein (Fragment).  
 GN KR.  
 OS Cupressus salei (Wandering spider).  
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;  
 OC Araneomorphae; Entelegynae; Lycosoidea; Ctenidae; Cupressius.  
 OX NCBI\_TaxID=628;  
 RN [1]  
 RP MEDLINE=93066327; PubMed=1438276;  
 RA Sommer R.J., Retzlaff M., Goerlich K., Sander K., Tautz D.;  
 RT "Evolutionary conservation pattern of zinc-finger domains of  
 Drosophila segmentation genes.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 89:10782-10786(1992).  
 CC -!- FUNCTION: KRUEPPEL IS A GAP CLASS SEGMENTATION PROTEIN.  
 CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).  
 CC -!- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-FINGER PROTEINS.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; L01594; AAA28292.1; -.  
 CC HSP; P08151; 2GLI.  
 CC InterPro: IPR000822; Znf-C2H2.  
 CC Pfam: PF00096; Zf-C2H2; 1.  
 CC SMART; SM00355; ZnF\_C2H2; 1.  
 CC PROSITE; PS00028; ZINC\_FINGER\_C2H2\_1; 1.  
 CC PROSITE; PS50157; ZINC\_FINGER\_C2H2\_2; 1.  
 KW Developmental protein; Gap protein; Zinc-finger;  
 KW Metal-binding; DNA-binding; Repeat; Nuclear protein.  
 FT NON\_TER 1 1  
 FT DOMAIN <1 >46 ZINC-FINGERS.  
 FT ZN\_FING <1 4 C2H2-TYPE.  
 FT ZN\_FING 10 32 C2H2-TYPE.  
 FT ZN\_FING 38 >46 C2H2-TYPE.  
 FT NON\_TER 46 46  
 SQ SEQUENCE 46 AA; 5546 MW; A92D4E6CECF2C1D2A CRC64;

Query Match 21.4%; Score 3; DB 1; Length 46;  
 Best Local Similarity 100.0%; Pred. No. 3.2e+02;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 HTG 5  
 Db 4 HTG 6

RESULT 12  
 KRUP\_LITFO  
 ID KRUP\_LITFO STANDARD; PRT; 46 AA.  
 AC Q01872;  
 DT 01-JUL-1993 (Rel. 26, Created)  
 DT 01-JUL-1993 (Rel. 26, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Krueppel protein (Fragment).  
 GN KR.  
 OS Lithobius forficatus.  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Myriapoda; Chilopoda;  
 OC Pleurostigmophora; Lithobiomorpha; Lithobiidae; Lithobius.  
 OX NCBI\_TaxID=7552;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93066327; PubMed=1438276;  
 RA Sommer R.J., Retzlaff M., Goerlich K., Sander K., Tautz D.;  
 RT "Evolutionary conservation pattern of zinc-finger domains of  
 Drosophila segmentation genes.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 89:10782-10786(1992).  
 CC -!- FUNCTION: KRUEPPEL IS A GAP CLASS SEGMENTATION PROTEIN.  
 CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).  
 CC -!- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-FINGER PROTEINS.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; L01598; AAA29280.1; -.  
 CC HSP; P08151; 2GLI.  
 CC InterPro: IPR000822; Znf-C2H2.  
 CC Pfam: PF00096; Zf-C2H2; 1.  
 CC SMART; SM00355; ZnF\_C2H2; 1.  
 CC PROSITE; PS00028; ZINC\_FINGER\_C2H2\_1; 1.  
 CC PROSITE; PS50157; ZINC\_FINGER\_C2H2\_2; 1.  
 KW Developmental protein; Gap protein; Zinc-finger;  
 KW Metal-binding; DNA-binding; Repeat; Nuclear protein.  
 FT NON\_TER 1 1  
 FT DOMAIN <1 >46 ZINC-FINGERS.  
 FT ZN\_FING <1 4 C2H2-TYPE.  
 FT ZN\_FING 10 32 C2H2-TYPE.  
 FT ZN\_FING 38 >46 C2H2-TYPE.  
 FT NON\_TER 46 46  
 SQ SEQUENCE 46 AA; 5493 MW; FEB4AF3297A72522 CRC64;

Query Match 21.4%; Score 3; DB 1; Length 46;  
 Best Local Similarity 100.0%; Pred. No. 3.2e+02;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 HTG 5  
 Db 4 HTG 6

RESULT 13  
 KRUP\_PHOPA  
 ID KRUP\_PHOPA STANDARD; PRT; 46 AA.  
 AC Q02034;

```
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Krueppel protein (Fragment).
GN KR.
OS Pholcus phalangioidea.
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;
OC Araneomorphae; Haplogynae; Pholcidae; Pholcus.
OX NCBI_TaxID=6352;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93066327; PubMed=1438276;
RA Sommer R.J., Retzlaff M., Goerlich K., Sander K., Tautz D.;
RT "Evolutionary conservation pattern of zinc-finger domains of
RT Drosophila segmentation genes."
RL Proc. Natl. Acad. Sci. U.S.A. 89:10782-10786(1992).
CC -!- FUNCTION: KRUEPPEL IS A GAP CLASS SEGMENTATION PROTEIN.
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-
CC FINGER PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L01605; AAA89211.1; -.
DR HSSP; P08151; 2GLI.
DR InterPro; IPR000822; Znf-C2H2.
DR Pfam; PF00096; zf-C2H2; 1.
DR SMART; SM00355; Znf.C2H2; 1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 1.
DR Developmental protein; Gap protein; Zinc-finger;
KW Metal-binding; DNA-binding; Repeat; Nuclear protein.
FT NON_TER 1
FT DOMAIN <1> >46 ZINC-FINGERS.
FT ZN_FING <1> 4 C2H2-TYPE.
FT ZN_FING 10 32 C2H2-TYPE.
FT ZN_FING 38 >46 C2H2-TYPE.
FT NON_TER 46 46
FT SEQUENCE 46 AA; 5515 MW; F323409A4A229D32 CRC64;

Query Match 21.4%; Score 3; DB 1; Length 46;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 HTG 5
Db 4 HTG 6

RESULT 14
LACA_LACAM STANDARD; PRT; 50 AA.
AC P80696;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE Bacteriocin lactobacin A.
OS Lactobacillus amylovorus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=1604;
RN [1]
RP SEQUENCE.
RC STRAIN=LMG P-13139;
RX MEDLINE=97133940; PubMed=8979334;
RA Contreras B.G.L., de Vuyst L., Devreese B., Busanyova K.,
```

```
RA Raymaekers J., Bosman F., Sablon E., Vandamme E.;
RT "Isolation, purification, and amino acid sequence of lactobacin A, one
RT of the two bacteriocins produced by Lactobacillus amylovorus LMG
RT P-13139."
RL Appl. Environ. Microbiol. 63:13-20(1997).
CC -!- FUNCTION: THIS HEAT STABLE BACTERIOCIN INHIBITS THE GROWTH OF
CC CLOSELY RELATED LACTOBACILLUS SPECIES. IT MAY ACT AS A PORE-
CC FORMING PROTEIN, CREATING A CHANNEL IN THE CELL MEMBRANE. IT KILLS
CC LACTOBACILLUS HELVETICUS ATCC 15009, BUT DISPLAYS NO ACTIVITY
CC TOWARDS LISTERIA SPECIES.
CC -!- SUBUNIT: ACTIVE LACTOBIN IS COMPOSED OF TWO DIFFERENT PEPTIDES,
CC ONE WHICH IS LACTOBIN A.
CC -!- MASS SPECTROMETRY: MW=4879; MW_ERR=0.69; METHOD=Electrospray.
CC -!- SIMILARITY: BELONGS TO THE CLASS IIB BACTERIOCINS.
KW Antibiotic; Bacteriocin; Transmembrane.
FT TRANSMEM 24
FT SEQUENCE 50 AA; 4882 MW; 663CD5F98CA03693 CRC64;

Query Match 21.4%; Score 3; DB 1; Length 50;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 GVM 9
Db 26 GVM 28

RESULT 15
ZN30_HUMAN STANDARD; PRT; 50 AA.
AC P17039;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Zinc finger protein 30 (Zinc finger protein KOX2s) (Fragment).
GN ZNF30 OR KOX28.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lymphoid;
RX MEDLINE=91145339; PubMed=2288909;
RA Thiesen H.-J.;
RT "Multiple genes encoding zinc finger domains are expressed in human T
RL cells."
RL New Biol. 2:363-374(1990).
CC -!- FUNCTION: MAY BE INVOLVED IN TRANSCRIPTIONAL REGULATION.
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-
CC FINGER PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X52359; CAA36585.1; -.
DR HSSP; P08151; 2GLI.
DR InterPro; IPR000822; Znf-C2H2.
DR Pfam; PF00096; zf-C2H2; 2.
DR SMART; SM00355; Znf.C2H2; 2.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 2.
KW Transcription regulation; Zinc-finger; DNA-binding; Metal-binding;
KW Nuclear protein; Repeat.
FT NON_TER 1
```



FT ZN\_FING 1 23 C2H2-TYPE.  
 FT ZN\_FING 29 >50 C2H2-TYPE.  
 FT NON\_TER 50 50  
 SQ SEQUENCE 50 AA; 5771 MW; 7786B30604FB3F00 CRC64;

Query Match 21.4%; Score 3; DB 1; Length 50;  
 Best Local Similarity 100.0%; Pred. No. 3.5e+02;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 HTG 5  
 |||  
 Db 23 HTG 25

Search completed: October 4, 2002, 10:36:36  
 Job time: 383 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 4, 2002, 10:36:14 ; Search time 43.48 Seconds  
(without alignments)  
55.702 Million cell updates/sec

Title: US-09-822-698A-28  
Perfect score: 14  
Sequence: 1 XXHTGXVWPXXX 14

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 562222 seqs, 172994929 residues

Word size : 0

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

SPTREMBL\_19:\*  
1: sp-archaea:\*  
2: sp-bacteria:\*  
3: sp-fungi:\*  
4: sp-human:\*  
5: sp-invertebrate:\*  
6: sp-mammal:\*  
7: sp-mhc:\*  
8: sp-organelle:\*  
9: sp-phage:\*  
10: sp-plant:\*  
11: sp-rodent:\*  
12: sp-virus:\*  
13: sp-vertebrate:\*  
14: sp-unclassified:\*  
15: sp-rvirus:\*  
16: sp-bacteriap:\*  
17: sp-archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4	28.6	60	10 Q41962	Q41962 arabidopsis
2	4	28.6	236	5 Q9Y1Q5	Q9Y1Q5 tridentator
3	4	28.6	349	10 Q9RSI7	Q9RSI7 hordeum vul
4	3	21.4	17	6 Q9XSG1	Q9XSG1 bos taurus
5	3	21.4	17	11 O88868	O88868 mus musculus
6	3	21.4	20	8 Q9MME9	Q9MME9 buteo hemil
7	3	21.4	22	4 Q9NNY7	Q9NNY7 homo sapien
8	3	21.4	22	11 Q9QV92	Q9QV92 rattus sp.
9	3	21.4	23	2 Q9R410	Q9R410 bordetella
10	3	21.4	23	2 Q93NC3	Q93NC3 myxococcus
11	3	21.4	23	5 Q9S895	Q9S895 drosophila
12	3	21.4	24	2 Q53392	Q53392 pseudomonas
13	3	21.4	24	5 Q77224	Q77224 helix asper
14	3	21.4	26	5 Q9BMD5	Q9BMD5 habrobrocha
15	3	21.4	27	11 P97646	P97646 rattus norv
16	3	21.4	27	15 Q85637	Q85637 rauscher sp

17	3	21.4	28	2	Q50102	Q50102 mycobacteri
18	3	21.4	28	11	Q60568	Q60568 mus cookii.
19	3	21.4	29	2	Q53393	Q53393 pseudomonas
20	3	21.4	29	6	Q9TSS7	Q9TSS7 felis silve
21	3	21.4	29	6	Q9BEC8	Q9BEC8 pongo pygma
22	3	21.4	29	11	O88213	O88213 mus musculus
23	3	21.4	29	11	O88214	O88214 mus musculus
24	3	21.4	29	11	O88215	O88215 mus musculus
25	3	21.4	29	11	O88217	O88217 mus musculus
26	3	21.4	29	11	O88218	O88218 mus musculus
27	3	21.4	29	11	O88219	O88219 mus musculus
28	3	21.4	29	11	O88220	O88220 mus musculus
29	3	21.4	29	11	O88221	O88221 mus musculus
30	3	21.4	29	11	O88222	O88222 mus musculus
31	3	21.4	29	11	O88223	O88223 mus musculus
32	3	21.4	29	11	O88224	O88224 mus musculus
33	3	21.4	29	11	O88225	O88225 mus musculus
34	3	21.4	29	11	O88227	O88227 mus musculus
35	3	21.4	29	11	O88228	O88228 mus musculus
36	3	21.4	29	11	Q9QWN1	Q9QWN1 mus musculus
37	3	21.4	29	11	Q9QWM6	Q9QWM6 mus musculus
38	3	21.4	29	11	Q9QWM5	Q9QWM5 mus musculus
39	3	21.4	32	6	Q9TSX3	Q9TSX3 equus cabal
40	3	21.4	32	6	Q9TSX0	Q9TSX0 sus scrofa
41	3	21.4	32	11	Q9QWN2	Q9QWN2 mus musculus
42	3	21.4	32	11	P97642	P97642 rattus norv
43	3	21.4	32	11	P97649	P97649 rattus norv
44	3	21.4	32	12	Q66861	Q66861 foot-and-mo
45	3	21.4	33	2	Q56122	Q56122 salmonella
			33	11	O88216	O88216 mus musculus

## ALIGNMENTS

RESULT 1

ID Q41962 PRELIMINARY; PRT; 60 AA.  
AC Q41962;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE FERRITIN (FRAGMENT).  
OS Arabidopsis thaliana (Mouse-ear cross).  
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidops.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. COLUMBIA; TISSUE=SEEDLING;  
RA Desprez T., Anselm J., Chiapello H., Rouze P., Caboche M., Hofte H.;  
RL Submitted (NOV-1992) to the EMBL/GenBank/DBJ databases.  
DR EMBL; Z18109; CAA79102.1; -;  
DR InterPro; IPR001519; Ferritin.  
DR Pfam; PF00210; ferritin; 1.  
DR ProDom; PD000971; Ferritin; 1.  
FT NON\_TER  
SQ SEQUENCE 60 AA; 7057 MW; B9849A05BB912F33 (R04);

Query Match 28.6%; Score 4; DB 10; Length 60;  
Best Local Similarity 100.0%; Pred. No. 66;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 GXGVW 9  
|||||  
DB 48 GXGVW 52

RESULT 2

Q9Y1Q5  
ID Q9Y1Q5 PRELIMINARY; PRT; 236 AA.  
AC Q9Y1Q5;

DT 01-NOV-1999 (TREMBLrel. 12, Created)  
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE DOPA DECARBOXYLASE (FRAGMENT).  
 OS Tridantiforma fuscoleuca.  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata;  
 OC Incurvarioidea; Prodoxidae; Tridantiforma.  
 OX NCBI\_TaxID=30245;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Friedlander T.P., Regier J.C., Mitter C., Wagner D.L.;  
 RT "Relationships within Heteroneuran Lepidoptera: Evidence from Dopa  
 RT Decarboxylase".  
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.  
 CC -/- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).  
 CC -/- SIMILARITY: BELONGS TO THE GNTR FAMILY OF TRANSCRIPTIONAL  
 CC REGULATORS.  
 CC -/- SIMILARITY: BELONGS TO GROUP II DECARBOXYLASES (DDC, GAD, HDC AND  
 CC TYRDC).  
 DR EMBL; AF093191; RAD42734.1; -.  
 DR InterPro; IPR000524; HTG\_GntR.  
 DR InterPro; IPR002129; Pyridoxal\_dec.  
 DR Pfam; PF00282; pyridoxal\_dec; 1.  
 DR PRINTS; PR00035; YHGNTR.  
 DR PRINTS; PR00800; YHDCRBOXLASE.  
 DR PROSITE; PS00352; DDC\_GAD\_HDC\_YDC; 1.  
 DR DNA-binding; decarboxylase; Lyase; Pyridoxal phosphate;  
 KW Transcription regulation.  
 FT NON\_TER 1 1  
 FT NON\_TER 236 236  
 SQ SEQUENCE 236 AA; 25865 MW; 58071346FED9CDB CRC64;

Query Match 28.6%; Score 4; DB 5; Length 236;  
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 TGXGV 8  
 Db 33 TGXGV 37

RESULT 3  
 ID Q9FSI7 PRELIMINARY; PRT; 349 AA.  
 AC Q9FSI7;  
 DT 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE PLASMA MEMBRANE PROTON ATPASE (FRAGMENT).  
 GN PPAL.  
 OS Hordeum vulgare (Barley).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;  
 OC Triticeae; Hordeum.  
 OX NCBI\_TaxID=4513;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=LEAF;  
 RA Deering-Saad C., Pritchard J., Newbury J., Bale J.;  
 RT "Detection of various mRNAs in barley phloem sap".  
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ295612; CAC10554.1; -.  
 DR InterPro; IPR001757; E1-E2\_ATPase.  
 DR InterPro; IPR000695; HATPase.  
 DR PRINTS; PR00119; CATATPASE.  
 DR PRINTS; PR00120; HATPASE.  
 FT NON\_TER 1 1  
 FT NON\_TER 349 349  
 SQ SEQUENCE 349 AA; 39172 MW; 896EF0FF7F03B136 CRC64;

Query Match 28.6%; Score 4; DB 10; Length 349;  
 Best Local Similarity 100.0%; Pred. No. 3.2e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 4 TGXGV 8  
 Db 6 TGXGV 10

## RESULT 4

ID Q9XSG1 PRELIMINARY; PRT; 17 AA.  
 AC Q9XSG1;  
 DT 01-NOV-1999 (TREMBLrel. 12, Created)  
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update);  
 DE SMCX (FRAGMENT).  
 GN SMCX.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=LIVER;  
 RA Polunienko A., Blecher S.;  
 RT "Exon-intron structure of SMCX and SMCY genes in bovine and swine".  
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF135448; AAD34440.1; -.  
 FT NON\_TER 1 1  
 FT NON\_TER 17 17  
 SQ SEQUENCE 17 AA; 1927 MW; 10351B0D516D16F0 CRC64;

Query Match 21.4%; Score 3; DB 6; Length 17;  
 Best Local Similarity 100.0%; Pred. No. 6.5e+02;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 GWV 9  
 Db 7 GWV 9

## RESULT 5

ID Q88868 PRELIMINARY; PRT; 17 AA.  
 AC Q88868;  
 DT 01-NOV-1998 (TREMBLrel. 08, Created)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE AMYLOID PROTEIN (FRAGMENT).  
 GN APP.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BALB/C; TISSUE=LIVER;  
 RX MEDLINE=20189613; PubMed=10727079;  
 RA Li L., Ohman T., Deeb S.S., Fukuchi K.-I.;  
 RT "Analysis of mouse intron 7 DNA sequence of the APP gene: comparison  
 RT with the human homologue".  
 RL DNA Seq. 10:219-228(1999).  
 DR EMBL; AF056079; AAC63987.1; -.  
 FT NON\_TER 1 1  
 FT NON\_TER 17 17  
 SQ SEQUENCE 17 AA; 1951 MW; 791C4AEB8464EADF CRC64;

Query Match 21.4%; Score 3; DB 11; Length 17;  
 Best Local Similarity 100.0%; Pred. No. 6.5e+02;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 GW 9  
|||  
DB 7 GW 9

## RESULT 6

Q9MME9 ID Q9MME9 PRELIMINARY; PRT; 20 AA.  
AC Q9MME9  
DT 01-OCT-2000 (TReMBLrel. 15, Created)  
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)  
DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)  
DE NADH DEHYDROGENASE SUBUNIT 6 (FRAGMENT).  
GN ND6.  
OS Buteo hemilasius (upland buzzard).  
OC Mitochondrion.  
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Falconiformes; Accipitridae;  
OC Accipitrinae; Buteo.  
OX NCBI\_TaxID=115122;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=HEM1;  
RA Haring E., Riesing M.J., Pinsker W., Gamauf A.;  
RT "Evolution of a pseudo-control region in the mitochondrial genome of  
RT Palearctic Buzzards (genus Buteo).";  
RL J. Zool. Syst. Evol. Res. 37:185-194 (1999).  
DR EMBL; AF202217; AAF61899.1; -.  
KW Mitochondrion.  
FT NON\_TER 20  
SQ SEQUENCE 20 AA; 2396 MW; 7B399CF466DA550C CRC64;

Query Match 21.4%; Score 3; DB 8; Length 20;

Best Local Similarity 100.0%; Pred. No. 7.5e+02;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 GW 9  
|||  
DB 15 GW 17

## RESULT 7

Q9NNY7 ID Q9NNY7 PRELIMINARY; PRT; 22 AA.  
AC Q9NNY7  
DT 01-OCT-2000 (TReMBLrel. 15, Created)  
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)  
DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)  
DE CYCLOOXYGENASE-2 (EC 1.14.99.1) (FRAGMENT).  
GN PTGS2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=EPIDERMAL KERATINOCYTES IN PRIMARY CULTURE;  
RX MEDLINE=98188304; PubMed=9520467;  
RA Lukiw W.J., Pelaez R.P., Martinez J., Bazan N.G.;  
RT "Budesonide epimer R or dexamethasone selectively inhibit platelet-  
RT activating factor-induced or interleukin beta-induced DNA binding  
RT activity of cis-acting transcription factors and cyclooxygenase-2 gene  
RT expression in human epidermal keratinocytes.";  
RL Proc. Natl. Acad. Sci. U.S.A. 95:3914-3919 (1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=EPIDERMAL KERATINOCYTES IN PRIMARY CULTURE;  
RX MEDLINE=98188304; PubMed=9520467;  
RA Lukiw W.J., Bazan N.G.;  
RT "Regulatory features of the human cyclooxygenase-2 (COX-2) promoter.";  
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF276953; AAF85977.1; -.  
KW Oxidoreductase.  
FT NON\_TER 22  
SQ SEQUENCE 22 AA; 2369 MW; A9B972BF5E5281E7 CRC64;

Query Match 21.4%; Score 3; DB 4; Length 22;

Best Local Similarity 100.0%; Pred. No. 8.2e+02;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 HTG 5  
|||  
DB 16 HTG 18

## RESULT 8

Q9QV92 ID Q9QV92 PRELIMINARY; PRT; 22 AA.  
AC Q9QV92  
DT 01-MAY-2000 (TReMBLrel. 13, Created)  
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)  
DE P93 CALCIUM-BINDING PROTEIN (FRAGMENT).  
OS Rattus sp.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Murinae; Rattus.  
OX NCBI\_TaxID=101118;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=93179436; PubMed=8440713;  
RA Gilchrist J.S., Pierce G.N.;  
RT "Identification and purification of a calcium-binding protein in  
RT hepatic nuclear membranes.";  
RL J. Biol. Chem. 268:4291-4299 (1993).  
DR InterPro; IPR001580; Calreticulin.  
DR Pfam; PF00262; Calreticulin; 1.  
DR ProDom; PD001866; Calreticulin; 1.  
SQ SEQUENCE 22 AA; 2407 MW; B82830E4477E25F9 CRC64;

Query Match 21.4%; Score 3; DB 11; Length 22;

Best Local Similarity 100.0%; Pred. No. 8.2e+02;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 GW 9  
|||  
DB 8 GW 10

## RESULT 9

Q9R410 ID Q9R410 PRELIMINARY; PRT; 23 AA.  
AC Q9R410  
DT 01-MAY-2000 (TReMBLrel. 13, Created)  
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)  
DE 32 KDA OUTER MEMBRANE PROTEIN (FRAGMENT).  
OS Bordetella pertussis.  
OC Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae;  
OC Bordetella.  
OX NCBI\_TaxID=520;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=96100814; PubMed=7483790;  
RA Hamstra H.J., Kuipers B., Schijff-Evers D., Loggen H.G., Poolman J.T.;  
RT "The purification and protective capacity of Bordetella pertussis  
RT outer membrane proteins.";  
RL Vaccine 13:747-752 (1995).  
SQ SEQUENCE 23 AA; 2473 MW; 0A9886CFDA308116 CRC64;

Query Match 21.4%; Score 3; DB 2; Length 23;

Best Local Similarity 100.0%; Pred. No. 8.5e+02;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 GWV 9  
|||  
Db 17 GWV 19

RESULT 10  
Q93NC3 PRELIMINARY; PRT; 23 AA.  
AC Q93NC3  
DT 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DE HYPOTHETICAL 2.5 KDA PROTEIN (FRAGMENT).  
OS Myxococcus xanthus.  
OC Bacteria; Proteobacteria; delta subdivision; Myxobacteria;  
OC Myxococcales; Cystobacterineae; Myxococcaceae; Myxococcus.  
OX NCBI\_TaxID=34;  
[1]  
RN SEQUENCE FROM N.A.  
RC STRAIN=DZFT;  
RA Nariya H., Inouye S.;  
RT "Identification of serine/threonine kinase associate proteins in M.  
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF377339; AAK64447.1; -;  
KW Hypothetical protein.  
FT NON\_TER 23  
SQ SEQUENCE 23 AA; 2517 MW; 71FACE29BAC9B335 CRC64;

Query Match 21.4%; Score 3; DB 2; Length 23;  
Best Local Similarity 100.0%; Pred. No. 8.5e+02;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 HTG 5  
|||  
Db 8 HTG 10

RESULT 11  
Q95S95 PRELIMINARY; PRT; 23 AA.  
AC Q95S95  
DT 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DE QM05003P.  
GN ANTP.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
[1]  
RN SEQUENCE FROM N.A.  
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,  
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,  
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,  
RA Nunoo J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,  
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;  
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AY060900; AAL28448.1; -;  
SQ SEQUENCE 23 AA; 2820 MW; 6F16D52A47F69139 CRC64;

Query Match 21.4%; Score 3; DB 5; Length 23;  
Best Local Similarity 100.0%; Pred. No. 8.5e+02;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 GWV 9  
|||

Db 8 GWV 10

RESULT 12  
Q53392 PRELIMINARY; PRT; 24 AA.  
ID Q53392  
AC Q53392;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE PIL PROTEIN (FRAGMENT).  
GN PIL.  
OS Pseudomonas aeruginosa.  
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
OC Pseudomonas.  
OX NCBI\_TaxID=287;  
[1]  
RN SEQUENCE FROM N.A.  
RP MEDLINE=94103636; PubMed=7903973;  
RX IPATSG;  
RA "A multicenter comparison of methods for typing strains of Pseudomonas  
RT aeruginosa predominantly from patients with cystic fibrosis. The  
RT International Pseudomonas aeruginosa Typing Study Group.";  
J. Infect. Dis. 169:134-142(1994).  
DR EMBL: S67811; CAB32863.1; -;  
FT NON\_TER 1  
SQ SEQUENCE 24 AA; 2583 MW; D2F8D4594FOCB689 CRC64;

Query Match 21.4%; Score 3; DB 2; Length 24;  
Best Local Similarity 100.0%; Pred. No. 8.8e+02;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 GWV 9  
|||  
Db 4 GWV 6

RESULT 13  
O77224 PRELIMINARY; PRT; 26 AA.  
ID O77224  
AC O77224;  
DT 01-NOV-1998 (TREMBLrel. 08, Created)  
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)  
DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)  
DE SHAKER-LIKE POTASSIUM CHANNEL KV1.1 (FRAGMENT).  
GN KV1.1.  
OS Helix aspersa (Brown garden snail).  
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Stylommatophora;  
OC Helicacea; Helicidae; Helix.  
OX NCBI\_TaxID=6535;  
[1]  
RN SEQUENCE FROM N.A.  
RP TISSUE=BRAIN;  
RC Witche H.J., Davies D., Meech R.W.;  
RA "Immunological identification of a putative potassium channel subunit  
RT from a short (9) amino acid sequence.";  
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF090122; AAC63039.1; -;  
KW Ionic channel.  
FT NON\_TER 1  
FT NON\_TER 26  
SQ SEQUENCE 26 AA; 2629 MW; 4500FCC00E3FF648 CRC64;

Query Match 21.4%; Score 3; DB 5; Length 26;  
Best Local Similarity 100.0%; Pred. No. 9.5e+02;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 GWV 9  
|||  
Db 6 GWV 8

```

RESULT 14
Q9BM05
AC Q9BM05 PRELIMINARY; PRT; 26 AA.
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE MARINER-LIKE TRANSPOSASE (FRAGMENT).
OS Habrotrocha constricta.
OC Eukaryota; Metazoa; Rotifera; Bdelloidea; Philodinida; Habrotrochidae;
OC Habrotrocha.
OX NCBI_TaxID=104786;
RN [1]
RP SEQUENCE FROM N.A.
RC TRANSPONSON=MARM4-6;
RX MEDLINE=20570504; PubMed=11121049;
RA Arkhipova I., Meselson M.;
RT "Transposable elements in sexual and ancient asexual taxa.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:14473-14477(2000).
DR EMBL; AY014001; AAG59973.1; -.
FT NON_TER 1
FT NON_TER 26
SQ SEQUENCE 26 AA; 2972 MW; 75D1BA812B90635C CRC64;

```

```

Query Match 21.4%; Score 3; DB 5; Length 26;
Best Local Similarity 100.0%; Pred. No. 9.5e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 HTG 5
Db 4 HTG 6

```

```

RESULT 15
P97646
ID P97646 PRELIMINARY; PRT; 27 AA.
AC P97646;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DE 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE ZINC FINGER PROTEIN 7 (FRAGMENT).
CN AZF7.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96220733; PubMed=8635150;
RA Mesa K., Gebelein B., Cook T., Urrutia R.;
RT "Identification and characterization of zinc finger encoding genes
from the tumoral exocrine pancreatic cell line AR42J.";
RL Cancer Lett. 103:143-149(1996).
DR EMBL; U78118; AAB36790.1; -.
FT NON_TER 1
FT NON_TER 27
SQ SEQUENCE 27 AA; 2750 MW; BDE7EF625ED05BBB CRC64;

```

```

Query Match 21.4%; Score 3; DB 11; Length 27;
Best Local Similarity 100.0%; Pred. No. 9.8e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 HTG 5
Db 21 HTG 23

```

Search completed: October 4, 2002, 10:36:15  
Job time: 387 sec

---



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run On: October 4, 2002, 10:21:47 ; Search time 51.14 seconds  
(without alignments)  
30.407 Million cell updates/sec

Title: US-09-822-698a-28  
Perfect score: 54  
Sequence: 1 XXHTGXGVWXPXXX 14

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_032802.\*

```

1: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	46	85.2	14	22	AAE12750
2	46	85.2	14	22	AAE12760
3	46	85.2	14	22	AAE12771
4	45	83.3	14	22	AAE12716
5	45	83.3	14	22	AAE12717
6	45	83.3	14	22	AAE12718
7	45	83.3	14	22	AAE12719
8	45	83.3	14	22	AAE12720
9	45	83.3	14	22	AAE12738
10	45	83.3	14	22	AAE12739
11	45	83.3	14	22	AAE12740

12	45	83.3	14	22	AAE12741	Human PH1 Fab anti
13	45	83.3	14	22	AAE12742	Human PH1 Fab anti
14	45	83.3	14	22	AAE12743	Human PH1 Fab anti
15	45	83.3	14	22	AAE12744	Human PH1 Fab anti
16	45	83.3	14	22	AAE12745	Human PH1 Fab anti
17	45	83.3	14	22	AAE12746	Human PH1 Fab anti
18	45	83.3	14	22	AAE12747	Human PH1 Fab anti
19	45	83.3	14	22	AAE12748	Human PH1 Fab anti
20	45	83.3	14	22	AAE12749	Human PH1 Fab anti
21	45	83.3	14	22	AAE12751	Human PH1 Fab anti
22	45	83.3	14	22	AAE12752	Human PH1 Fab anti
23	45	83.3	14	22	AAE12753	Human PH1 Fab anti
24	45	83.3	14	22	AAE12754	Human PH1 Fab anti
25	45	83.3	14	22	AAE12755	Human PH1 Fab anti
26	45	83.3	14	22	AAE12756	Human PH1 Fab anti
27	45	83.3	14	22	AAE12757	Human PH1 Fab anti
28	45	83.3	14	22	AAE12758	Human PH1 Fab anti
29	45	83.3	14	22	AAE12759	Human PH1 Fab anti
30	45	83.3	14	22	AAE12761	Human PH1 Fab anti
31	45	83.3	14	22	AAE12762	Human PH1 Fab anti
32	45	83.3	14	22	AAE12763	Human PH1 Fab anti
33	45	83.3	14	22	AAE12764	Human PH1 Fab anti
34	45	83.3	14	22	AAE12765	Human PH1 Fab anti
35	45	83.3	14	22	AAE12766	Human PH1 Fab anti
36	45	83.3	14	22	AAE12767	Human PH1 Fab anti
37	45	83.3	14	22	AAE12768	Human PH1 Fab anti
38	45	83.3	14	22	AAE12769	Human PH1 Fab anti
39	45	83.3	14	22	AAE12770	Human PH1 Fab anti
40	45	83.3	14	22	AAE12772	Human PH1 Fab anti
41	45	83.3	14	22	AAE12773	Human PH1 Fab anti
42	45	83.3	14	22	AAE12774	Human PH1 Fab anti
43	45	83.3	16	22	AAE12735	Human PH1 Fab anti
44	45	83.3	16	22	AAE12737	Human PH1 Fab anti
45	45	83.3	121	22	AAE12706	Human PH1 Fab anti

#### ALIGNMENTS

RESULT 1

AAE12750

ID AAE12750 standard; peptide; 14 AA.

XX AAE12750;

XX 04-JAN-2002 (first entry)

DE Human PH1 Fab antibody VH region PR3-CDR3 variant #17.

XX Human; tumour-associated antigen mucin-1; MUC-1; adenocarcinoma; VH;  
KW variable heavy chain region; cancer; breast; ovary; lung; bladder;  
KW cytostatic; therapy; PH1 antibody; variant.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Region 1..2

FT Region /label= Framework\_region\_3

FT Region 3..14

FT /label= Complementarity\_determining\_region\_3

XX WO200175110-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US10589.

XX 30-MAR-2000; 2000US-0538913.

XX (DYAX-) DYAX CORP.

XX Hoogenboom HRJM, Henderikx MPG;

XX

DR WPI; 2001-626437/72.

XX Novel isolated tumor-associated antigen mucin-1-specific binding member

PT for diagnosing and treating cancer, comprises mucin-1 binding domain or

PT its portion for binding to an epitope of the protein core of mucin-1

XX

PS Example 2; Page 121; 126pp; English.

XX The invention relates to an isolated tumour-associated antigen mucin-1

CC (MUC-1)-specific binding member comprising an antigen binding domain

CC region having an antibody variable light (VL) or heavy (VH) region,

CC or a complementarity determining region (CDR) of VL or VH. MUC1-specific

CC binding member is useful for diagnosing cancer, preferably adenocarcinoma

CC The binding of MUC1-specific binding member to MUC1 is detected by a

CC detection method selected from enzyme-linked immunosorbent assay,

CC magnetic resonance imaging, scintillation counting, and X-ray film.

CC MUC1-specific binding member is useful for treating cancer, preferably

CC adenocarcinoma, in an individual, where the cancer is present in tissue

CC of the breast, ovary, lung, or bladder of the individual. MUC1-specific

CC binding member is useful for diagnosing and imaging MUC1-expressing

CC cancer cells and tissues, for purifying or isolating non-glycosylated,

CC underglycosylated or cancer-associated forms of MUC1 or MUC1 epitope-

CC containing molecules, and for therapeutically or prophylactically

CC treating cancer. The present sequence is human PH1 Fab antibody VH

CC region FR (framework region)3-CDR3 variant.

XX

SQ Sequence 14 AA;

Query Match 85.2%; Score 46; DB 22; Length 14;

Best Local Similarity 77.8%; Pred. No. 0.16;

Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 3 HTGXGVWXP 11

Db 111 111 1

3 htgsgvwdp 11

RESULT 2

AAE12760

ID AAE12760 standard; peptide; 14 AA.

XX

AC AAE12760;

XX

XX 04-JAN-2002 (first entry)

DE Human PH1 Fab antibody VH region FR3-CDR3 variant. #27.

XX Human; tumour-associated antigen mucin-1; MUC-1; adenocarcinoma; VH;

KW variable heavy chain region; cancer; breast; ovary; lung; bladder;

KW cytostatic; therapy; PH1 antibody; variant.

XX

OS Homo sapiens.

XX

PH Key Location/Qualifiers

FT Region 1..2

FT /label= Framework\_region\_3

FT Region 3..14

FT /label= Complementarity\_determining\_region\_3

XX

XX WO200175110-A2.

XX

XX 11-OCT-2001.

XX

XX 30-MAR-2001; 2001WO-US10589.

XX

XX 30-MAR-2000; 2000US-0538913.

XX

XX (DYAX-) DYAX CORP.

XX

XX Hoogenboom HRJM, Henderikx MPG;

XX

DR WPI; 2001-626437/72.

XX Novel isolated tumor-associated antigen mucin-1-specific binding member

PT for diagnosing and treating cancer, comprises mucin-1 binding domain or

PT its portion for binding to an epitope of the protein core of mucin-1

XX

PS Example 2; Page 123; 126pp; English.

XX The invention relates to an isolated tumour-associated antigen mucin-1

CC (MUC-1)-specific binding member comprising an antigen binding domain

CC region having an antibody variable light (VL) or heavy (VH) region,

CC or a complementarity determining region (CDR) of VL or VH. MUC1-specific

CC binding member is useful for diagnosing cancer, preferably adenocarcinoma

CC The binding of MUC1-specific binding member to MUC1 is detected by a

CC detection method selected from enzyme-linked immunosorbent assay,

CC magnetic resonance imaging, scintillation counting, and X-ray film.

CC MUC1-specific binding member is useful for treating cancer, preferably

CC adenocarcinoma, in an individual, where the cancer is present in tissue

CC of the breast, ovary, lung, or bladder of the individual. MUC1-specific

CC binding member is useful for diagnosing and imaging MUC1-expressing

CC cancer cells and tissues, for purifying or isolating non-glycosylated,

CC underglycosylated or cancer-associated forms of MUC1 or MUC1 epitope-

CC containing molecules, and for therapeutically or prophylactically

CC treating cancer. The present sequence is human PH1 Fab antibody VH

CC region FR (framework region)3-CDR3 variant.

XX

SQ Sequence 14 AA;

Query Match 85.2%; Score 46; DB 22; Length 14;

Best Local Similarity 77.8%; Pred. No. 0.16;

Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 3 HTGXGVWXP 11

Db 111 111 1

3 htgsgvwdp 11

RESULT 3

AAE12771

ID AAE12771 standard; peptide; 14 AA.

XX

AC AAE12771;

XX

XX 04-JAN-2002 (first entry)

DE Human PH1 Fab antibody VH region FR3-CDR3 variant. #18.

XX Human; tumour-associated antigen mucin-1; MUC-1; adenocarcinoma; VH;

KW variable heavy chain region; cancer; breast; ovary; lung; bladder;

KW cytostatic; therapy; PH1 antibody; variant.

XX

OS Homo sapiens.

XX

PH Key Location/Qualifiers

FT Region 1..2

FT /label= Framework\_region\_3

FT Region 3..14

FT /label= Complementarity\_determining\_region\_3

XX

XX WO200175110-A2.

XX

XX 11-OCT-2001.

XX

XX 30-MAR-2001; 2001WO-US10589.

XX

XX 30-MAR-2000; 2000US-0538913.

XX

XX (DYAX-) DYAX CORP.

XX

XX Hoogenboom HRJM, Henderikx MPG;

XX

DR WPI; 2001-626437/72.

PT Novel isolated tumor-associated antigen mucin-1-specific binding member  
PT for diagnosing and treating cancer, comprises mucin-1 binding domain or  
PT its portion for binding to an epitope of the protein core of mucin-1 -  
XX  
XX  
XX Example 2; Page 125; 126pp: English.

XX The invention relates to an isolated tumour-associated antigen mucin-1  
CC (MUC-1)-specific binding member comprising an antigen binding domain  
CC region having an antibody variable light (VL) or heavy (VH) region,  
CC or a complementarity determining region (CDR) of VL or VH. MUC1-specific  
CC binding member is useful for diagnosing cancer, preferably adenocarcinoma  
CC The binding of MUC1-specific binding member to MUC1 is detected by a  
CC detection method selected from enzyme-linked immunosorbent assay,  
CC magnetic resonance imaging, scintillation counting, and X-ray film.  
CC MUC1-specific binding member is useful for treating cancer, preferably  
CC of the breast, ovary, lung, or bladder of the individual. MUC1-specific  
CC binding member is useful for diagnosing and imaging MUC1-expressing  
CC cancer cells and tissues, for purifying or isolating non-glycosylated,  
CC underglycosylated or cancer-associated forms of MUC1 or MUC1 epitope-  
CC containing molecules, and for therapeutically or prophylactically  
CC treating cancer. The present sequence is human PH1 Fab antibody VH  
CC region FR (framework region)3-CDR3 variant.  
XX  
XX Sequence 14 AA;

Query Match 85.2%; Score 46; DB 22; Length 14;  
Best Local Similarity 77.8%; Pred. No. 0.16;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 HTGCGVWXP 11  
| | | | |  
Db 3 htgagvwdp 11

RESULT 4  
AAE12716

ID AAE12716 standard; peptide; 14 AA.

AC AAE12716;

DT 04-JAN-2002 (first entry)

XX MUC1-specific binding member antigen binding domain.

XX Tumour-associated antigen mucin-1; MUC-1; adenocarcinoma; cytostatic;  
KW cancer; breast; ovary; lung; bladder; therapy.

XX Unidentified.

XX Key Location/Qualifiers

FT Misc-difference 1 /label= Ala, Ser, Thr, Val

FT Misc-difference 2 /label= Lys, Ile, Arg, Gln

FT Misc-difference 6 /label= Gly, Arg, Val, Glu, Ser, Ala

FT Misc-difference 10 /label= Asn, Asp

FT Misc-difference 12 /label= Ile, Leu, Met, Phe, Val

FT Misc-difference 13 /label= Asp, Gly, Lys, Asn, Ala, His, Arg, Ser, Val, Tyr

FT Misc-difference 14 /label= Tyr, His, Lys, Asn, Asp, Ser, Pro

XX WO200175110-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US10589.

XX

PR 30-MAR-2000; 2000US-0538913.  
XX (DYAX-) DYAX CORP.

XX Hoogenboom HRJM, Henderikx MPG;

XX WPI; 2001-626437/72.

XX Novel isolated tumor-associated antigen mucin-1-specific binding member  
PT for diagnosing and treating cancer, comprises mucin-1 binding domain or  
PT its portion for binding to an epitope of the protein core of mucin-1 -  
XX  
XX Claim 1; Page 75; 126pp: English.

XX The invention relates to an isolated tumour-associated antigen mucin-1  
CC (MUC-1)-specific binding member comprising an antigen binding domain  
CC region having an antibody variable light (VL) or heavy (VH) region,  
CC or a complementarity determining region (CDR) of VL or VH. MUC1-specific  
CC binding member is useful for diagnosing cancer, preferably adenocarcinoma  
CC The binding of MUC1-specific binding member to MUC1 is detected by a  
CC detection method selected from enzyme-linked immunosorbent assay,  
CC magnetic resonance imaging, scintillation counting, and X-ray film.  
CC MUC1-specific binding member is useful for treating cancer, preferably  
CC of the breast, ovary, lung, or bladder of the individual. MUC1-specific  
CC binding member is useful for diagnosing and imaging MUC1-expressing  
CC cancer cells and tissues, for purifying or isolating non-glycosylated,  
CC underglycosylated or cancer-associated forms of MUC1 or MUC1 epitope-  
CC containing molecules, and for therapeutically or prophylactically  
CC treating cancer. The present sequence is antigen binding domain of  
CC MUC1-specific binding member.  
XX  
XX Sequence 14 AA;

Query Match 83.3%; Score 45; DB 22; Length 14;  
Best Local Similarity 100.0%; Pred. No. 0.24;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 HTGCGVWXP 11  
| | | | |  
Db 3 htgagvwdp 11

RESULT 5

AAE12717

ID AAE12717 standard; peptide; 14 AA.

AC AAE12717;

DT 04-JAN-2002 (first entry)

XX Human PH1 Fab antibody VH region FR3-CDR3 variant #1.

XX Human; tumour-associated antigen mucin-1; MUC-1; adenocarcinoma; VH;  
KW variable heavy chain region; cancer; breast; ovary; lung; bladder;  
KW cytostatic; therapy; PH1 antibody; variant.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Region 1..2 /label= Framework\_region\_3

FT Region 3..14 /label= Complementarity\_determining\_region\_3

XX WO200175110-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US10589.

XX 30-MAR-2000; 2000US-0538913.

XX

XX (DYAX-) DYAX CORP.  
 XX Hoogenboom HRJM, Henderikx MPG;  
 XX WPI; 2001-626437/72.  
 XX Novel isolated tumor-associated antigen mucin-1-specific binding member  
 PT for diagnosing and treating cancer, comprises mucin-1 binding domain or  
 PT its portion for binding to an epitope of the protein core of mucin-1  
 XX  
 XX Claim 2; Page 75; 126pp; English.  
 XX The invention relates to an isolated tumour-associated antigen mucin-1  
 CC (MUC-1)-specific binding member comprising an antigen binding domain  
 CC region having an antibody variable light (VL) or heavy (VH) region,  
 CC or a complementarity determining region (CDR) of VL or VH. MUC1-specific  
 CC binding member is useful for diagnosing cancer, preferably adenocarcinoma  
 CC The binding of MUC1-specific binding member to MUC1 is detected by a  
 CC detection method selected from enzyme-linked immunosorbent assay,  
 CC magnetic resonance imaging, scintillation counting, and x-ray film.  
 CC MUC1-specific binding member is useful for treating cancer, preferably  
 CC adenocarcinoma, in an individual, where the cancer is present in tissue  
 CC of the breast, ovary, lung, or bladder of the individual. MUC1-specific  
 CC binding member is useful for diagnosing and imaging MUC1-expressing  
 CC cancer cells and tissues, for purifying or isolating non-glycosylated,  
 CC underglycosylated or cancer-associated forms of MUC1 or MUC1 epitope-  
 CC containing molecules, and for therapeutically or prophylactically  
 CC treating cancer. The present sequence is human PH1 Fab antibody VH  
 CC region FR (framework region)3-CDR3 variant.  
 XX Sequence 14 AA;  
 SQ

Query Match 83.3%; Score 45; DB 22; Length 14;  
 Best Local Similarity 77.8%; Pred. No. 0.24;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 3 HTGXGWWXP 11  
 DB 3 htgrgvwdp 11

RESULT 6  
 AAE12718  
 ID AAE12718 standard; peptide; 14 AA.  
 XX  
 AC AAE12718;  
 DT 04-JAN-2002 (first entry)  
 DT Human PH1 Fab antibody VH region FR3-CDR3 variant #2.  
 DE  
 DE Human; tumour-associated antigen mucin-1; MUC-1; adenocarcinoma; VH;  
 KW variable heavy chain region; cancer; breast; ovary; lung; bladder;  
 KW cytostatic; therapy; PH1 antibody; variant.  
 XX Homo sapiens.  
 XX  
 XX Key Location/Qualifiers  
 FH Region 1..2  
 FT /label= Framework\_region\_3  
 FT 3..14  
 FT /label= Complementarity\_determining\_region\_3  
 FT  
 FT WO200175110-A2.  
 PN 11-OCT-2001.  
 PD 30-MAR-2001; 2001WO-US10589.  
 XX 30-MAR-2000; 2000US-0538913.  
 XX (DYAX-) DYAX CORP.

PA (DYAX-) DYAX CORP.  
 XX Hoogenboom HRJM, Henderikx MPG;  
 XX WPI; 2001-626437/72.  
 XX Novel isolated tumor-associated antigen mucin-1-specific binding member  
 PT for diagnosing and treating cancer, comprises mucin-1 binding domain or  
 PT its portion for binding to an epitope of the protein core of mucin-1  
 XX  
 XX Claim 2; Page 75; 126pp; English.  
 XX The invention relates to an isolated tumour-associated antigen mucin-1  
 CC (MUC-1)-specific binding member comprising an antigen binding domain  
 CC region having an antibody variable light (VL) or heavy (VH) region,  
 CC or a complementarity determining region (CDR) of VL or VH. MUC1-specific  
 CC binding member is useful for diagnosing cancer, preferably adenocarcinoma  
 CC The binding of MUC1-specific binding member to MUC1 is detected by a  
 CC detection method selected from enzyme-linked immunosorbent assay,  
 CC magnetic resonance imaging, scintillation counting, and x-ray film.  
 CC MUC1-specific binding member is useful for treating cancer, preferably  
 CC adenocarcinoma, in an individual, where the cancer is present in tissue  
 CC of the breast, ovary, lung, or bladder of the individual. MUC1-specific  
 CC binding member is useful for diagnosing and imaging MUC1-expressing  
 CC cancer cells and tissues, for purifying or isolating non-glycosylated,  
 CC underglycosylated or cancer-associated forms of MUC1 or MUC1 epitope-  
 CC containing molecules, and for therapeutically or prophylactically  
 CC treating cancer. The present sequence is human PH1 Fab antibody VH  
 CC region FR (framework region)3-CDR3 variant.  
 XX Sequence 14 AA;  
 SQ

Query Match 83.3%; Score 45; DB 22; Length 14;  
 Best Local Similarity 77.8%; Pred. No. 0.24;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 3 HTGXGWWXP 11  
 DB 3 htgrgvwdp 11

RESULT 7  
 AAE12719  
 ID AAE12719 standard; peptide; 14 AA.  
 XX  
 AC AAE12719;  
 DT 04-JAN-2002 (first entry)  
 DT Human PH1 Fab antibody VH region FR3-CDR3 variant #3.  
 DE  
 DE Human; tumour-associated antigen mucin-1; MUC-1; adenocarcinoma; VH;  
 KW variable heavy chain region; cancer; breast; ovary; lung; bladder;  
 KW cytostatic; therapy; PH1 antibody; variant.  
 XX Homo sapiens.  
 XX  
 XX Key Location/Qualifiers  
 FH Region 1..2  
 FT /label= Framework\_region\_3  
 FT 3..14  
 FT /label= Complementarity\_determining\_region\_3  
 FT  
 FT WO200175110-A2.  
 PN 11-OCT-2001.  
 PD 30-MAR-2001; 2001WO-US10589.  
 XX 30-MAR-2000; 2000US-0538913.  
 XX (DYAX-) DYAX CORP.

XX	
XX	Hoogenboom HRJM, Henderikx MPG;
PI	
XX	WPI; 2001-626437/72.
DR	
XX	
XX	
XX	Novel isolated tumor-associated antigen mucin-1-specific binding member
PT	for diagnosing and treating cancer, comprises mucin-1 binding domain or
PT	its portion for binding to an epitope of the protein core of mucin-1
PT	-
PT	
PS	Claim 2; Page 75; 126pp; English.
XX	
XX	The invention relates to an isolated tumour-associated antigen mucin-1
CC	(MUC-1)-specific binding member comprising an antigen binding domain
CC	region having an antibody variable light (VL) or heavy (VH) region,
CC	or a complementarity determining region (CDR) of VL or VH. MUC1-specific
CC	binding member is useful for diagnosing cancer, preferably adenocarcinoma
CC	binding member is useful for diagnosing cancer, preferably adenocarcinoma
CC	The binding of MUC1-specific binding member to MUC1 is detected by a
CC	detection method selected from enzyme-linked immunosorbent assay,
CC	magnetic resonance imaging, scintillation counting, and X-ray film.
CC	MUC1-specific binding member is useful for treating cancer, preferably
CC	adenocarcinoma, in an individual, where the cancer is present in tissue
CC	of the breast, ovary, lung, or bladder of the individual. MUC1-specific
CC	binding member is useful for diagnosing and imaging MUC1-expressing
CC	cancer cells and tissues, for purifying or isolating non-glycosylated,
CC	cancer cells and tissues, for purifying or isolating non-glycosylated,
CC	underglycosylated or cancer-associated forms of MUC1 or MUC1 epitope-
CC	containing molecules, and for therapeutically or prophylactically
CC	treating cancer. The present sequence is human PHI Fab antibody VH
CC	region FR (framework region)3-CDR3 variant.
XX	
XX	Sequence 14 AA;
SQ	

Query Match	83.3%	Score 45;	DB 22;	Length 14;
Best Local Similarity	77.8%	Pred. No. 0.24;		
Matches	7;	Conservative	0;	Mismatches 2; Indels 0; Gaps
QY	3 HTGXGVMXP 11			
DB	3 htgggwdp 11			
RESULT	9			
AAE12738				
ID	AAE12738 standard; peptide; 14 AA.			
XX				
AC	AAE12738;			
XX				
DT	04-JAN-2002 (first entry)			
XX				
DE	Human PH1 Fab antibody VH region FR3-CDR3 variant #5.			
XX				
KW	Human; tumour-associated antigen mucin-1; MUC-1; adenocarcinoma; VH;			
KW	variable heavy chain region; cancer; breast; ovary; lung; bladder;			
KW	cytostatic; therapy; PH1 antibody; variant.			
XX				
OS	Homo sapiens.			
XX				
Key	Location/Qualifiers			
FT	Region 1..2			
FT	/label= Framework_region_3			
FT	Region 3..14			
FT	/label= Complementarity_determining_region_3			
XX				
PN	WO200175110-A2.			
XX				
PD	11-OCT-2001.			
XX				
PF	30-MAR-2001; 2001WO-US10589.			
XX				
PR	30-MAR-2000; 2000US-0538913.			
XX				
PA	(DYAX-) DYAX CORP.			
XX				
PI	Hoogenboom HRJM, Hendrikx MPG;			
XX				

```

PI Hoogenboom HRJM, Henderikx MPG;
XX
XX WPI: 2001-626437/72.
XX
XX Novel isolated tumor-associated antigen mucin-1-specific binding member
XX PT for diagnosing and treating cancer, comprises mucin-1 binding domain or
XX PT its portion for binding to an epitope of the protein core of mucin-1
XX
XX Claim 2; Page 75; 126pp; English.
XX
XX The invention relates to an isolated tumour-associated antigen mucin-1
XX CC (MUC-1)-specific binding member comprising an antigen binding domain
XX CC region having an antibody variable light (VL) or heavy (VH) region,
XX CC or a complementarity determining region (CDR) of VL or VH, MUC1-specific
XX CC binding member is useful for diagnosing cancer, preferably adenocarcinoma
XX CC The binding of MUC1-specific binding member to MUC1 is detected by a
XX CC detection method selected from enzyme-linked immunosorbent assay,
XX CC magnetic resonance imaging, scintillation counting, and X-ray film.
XX CC MUC1-specific binding member is useful for treating cancer, preferably
XX CC adenocarcinoma, in an individual, where the cancer is present in tissue
XX CC of the breast, ovary, lung, or bladder of the individual. MUC1-specific
XX CC binding member is useful for diagnosing and imaging MUC1-expressing
XX CC cancer cells and tissues, for purifying or isolating non-glycosylated,
XX CC underglycosylated or cancer-associated forms of MUC1 or MUC1 epitope-
XX CC containing molecules, and for therapeutically or prophylactically
XX CC treating cancer. The present sequence is human PH1 Fab antibody VH
XX CC region FR (framework region)3-CDR3 variant.
XX
XX Sequence 14 AA;
SQ
Query Match 83.3%; Score 45; DB 22; Length 14;
Best Local Similarity 77.8%; Pred. No. 0.24;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps
QY 3 HTGXGVWXP 11
Db ||| ||| |
3 htgggvgwdp 11
RESULT 9
AAE12738
ID AAE12738 standard; peptide; 14 AA.
AC XX
AC AAE12738;
XX
XX 04-JAN-2002 (first entry)
XX Human PH1 Fab antibody VH region FR3-CDR3 variant #5.
XX
XX Human; tumour-associated antigen mucin-1; MUC-1; adenocarcinoma; VH;
KW variable heavy chain region; cancer; breast; ovary; lung; bladder;
KW cytostatic; therapy; PH1 antibody; variant.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
XX FH 1..2
XX FT Region /label= Framework_region_3
XX FT Region 3..14
XX FT /label= Complementarity_determining_region_3
XX
XX WO200175110-A2.
XX
XX 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-0510589.
XX
XX 30-MAR-2000; 2000US-0538913.
XX
XX (DYAX-) DYAX CORP.
XX
XX Hoogenboom HRJM, Henderikx MPG;
PI

```

```

Query Match      83.3%; Score 45; DB 22; Length 14;
Best Local Similarity 77.8%; Pred. No. 0.24;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps

```

QY 3 HTGXGVWXP 11  
 111 11111

DB 3 htgggrwdp 11

RESULT 9  
 AAEL12738  
 ID AAE12738 standard; peptide; 14 AA.  
 XX XX  
 AC AAE12738;  
 XX XX  
 DT 04-JAN-2002 (first entry)  
 XX XX  
 DE Human PH1 Fab antibody VH region FR3-CDR3 variant #5.  
 XX XX  
 KW Human; tumour-associated antigen mucin-1; MUC-1; adenocarcinoma; VH;  
 KW variable heavy chain region; cancer; breast; ovary; lung; bladder;  
 KW cystostatic; therapy; PH1 antibody; variant.  
 XX XX  
 OS Homo sapiens.  
 XX XX  
 Key Location/Qualifiers  
 FH Region 1..2  
 FT Region /label= Framework\_region\_3  
 FT Region 3..14  
 FT Region /label= Complementarity\_determining\_region\_3  
 XX XX  
 WO200175110-A2.  
 XX XX  
 11-OCT-2001.  
 XX XX  
 30-MAR-2001; 2001WO-0510589.  
 XX XX  
 30-MAR-2000; 2000US-0538913.  
 PR XX  
 XX XX  
 (DYAX-) DYAX CORP.  
 PA XX  
 XX XX  
 Hoogenboom HRJM, Henderikx MPG;

XX WPI; 2001-626437/72.

XX Novel isolated tumor-associated antigen mucin-1-specific binding member

PT for diagnosing and treating cancer, comprises mucin-1 binding domain or

PT its portion for binding to an epitope of the protein core of mucin-1 -

XX

XX Example 2; Page 118; 126pp; English.

XX The invention relates to an isolated tumour-associated antigen mucin-1

CC (MUC-1)-specific binding member comprising an antigen binding domain

CC region having an antibody variable light (VL) or heavy (VH) region,

CC or a complementarity determining region (CDR) of VL or VH. MUC1-specific

CC binding member is useful for diagnosing cancer, preferably adenocarcinoma

CC The binding of MUC1-specific binding member to MUC1 is detected by a

CC detection method selected from enzyme-linked immunosorbent assay,

CC magnetic resonance imaging, scintillation counting, and X-ray film.

CC MUC1-specific binding member is useful for treating cancer, preferably

CC adenocarcinoma, in an individual, where the cancer is present in tissue

CC of the breast, ovary, lung, or bladder of the individual. MUC1-specific

CC binding member is useful for diagnosing and imaging MUC1-expressing

CC cancer cells and tissues, for purifying or isolating non-glycosylated,

CC underglycosylated or cancer-associated forms of MUC1 or MUC1 epitope-

CC containing molecules, and for therapeutically or prophylactically

CC treating cancer. The present sequence is human PH1 Fab antibody VH

CC region FR (framework region)3-CDR3 variant.

XX

XX Sequence 14 AA;

Query Match 83.38; Score 45; DB 22; Length 14;

Best Local Similarity 77.8%; Pred. No. 0.24;

Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 HTGXGVWXP 11

||| ||| |

Db 3 htgvgvwdp 11

RESULT 10

AAE12739

ID AAE12739 standard; peptide; 14 AA.

XX

XX AAE12739;

AC

XX 04-JAN-2002 (first entry)

DT

XX Human PH1 Fab antibody VH region FR3-CDR3 variant #6.

DE

XX Human; tumour-associated antigen mucin-1; MUC-1; adenocarcinoma; VH;

KW variable heavy chain region; cancer; breast; ovary; lung; bladder;

KW cytostatic; therapy; PH1 antibody; variant.

XX

OS Homo sapiens.

XX

XX Key Location/Qualifiers

FT Region 1..2

FT /label= Framework\_region\_3

FT Region 3..14

FT /label= Complementarity\_determining\_region\_3

FT

XX WO200175110-A2.

PN

XX 11-OCT-2001.

PD

XX 30-MAR-2001; 2001WO-US10589.

XX

XX 30-MAR-2000; 2000US-0538913.

XX

XX (DYAX-) DYAX CORP.

XX

XX Hoogenboom HRJM, Henderikx MPG;

XX

XX WPI; 2001-626437/72.

DR WPI; 2001-626437/72.

XX

XX Novel isolated tumor-associated antigen mucin-1-specific binding member

PT for diagnosing and treating cancer, comprises mucin-1 binding domain or

PT its portion for binding to an epitope of the protein core of mucin-1 -

XX

XX Example 2; Page 118; 126pp; English.

XX The invention relates to an isolated tumour-associated antigen mucin-1

CC (MUC-1)-specific binding member comprising an antigen binding domain

CC region having an antibody variable light (VL) or heavy (VH) region,

CC or a complementarity determining region (CDR) of VL or VH. MUC1-specific

CC binding member is useful for diagnosing cancer, preferably adenocarcinoma

CC The binding of MUC1-specific binding member to MUC1 is detected by a

CC detection method selected from enzyme-linked immunosorbent assay,

CC magnetic resonance imaging, scintillation counting, and X-ray film.

CC MUC1-specific binding member is useful for treating cancer, preferably

CC adenocarcinoma, in an individual, where the cancer is present in tissue

CC of the breast, ovary, lung, or bladder of the individual. MUC1-specific

CC binding member is useful for diagnosing and imaging MUC1-expressing

CC cancer cells and tissues, for purifying or isolating non-glycosylated,

CC underglycosylated or cancer-associated forms of MUC1 or MUC1 epitope-

CC containing molecules, and for therapeutically or prophylactically

CC treating cancer. The present sequence is human PH1 Fab antibody VH

CC region FR (framework region)3-CDR3 variant.

XX

XX Sequence 14 AA;

Query Match 83.38; Score 45; DB 22; Length 14;

Best Local Similarity 77.8%; Pred. No. 0.24;

Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 HTGXGVWXP 11

||| ||| |

Db 3 htgvgvwdp 11

RESULT 11

AAE12740

ID AAE12740 standard; peptide; 14 AA.

XX

XX AAE12740;

AC

XX 04-JAN-2002 (first entry)

DT

XX Human PH1 Fab antibody VH region FR3-CDR3 variant #7.

DE

XX Human; tumour-associated antigen mucin-1; MUC-1; adenocarcinoma; VH;

KW variable heavy chain region; cancer; breast; ovary; lung; bladder;

KW cytostatic; therapy; PH1 antibody; variant.

XX

OS Homo sapiens.

XX

XX Key Location/Qualifiers

FT Region 1..2

FT /label= Framework\_region\_3

FT Region 3..14

FT /label= Complementarity\_determining\_region\_3

FT

XX WO200175110-A2.

PN

XX 11-OCT-2001.

PD

XX 30-MAR-2001; 2001WO-US10589.

XX

XX 30-MAR-2000; 2000US-0538913.

XX

XX (DYAX-) DYAX CORP.

XX

XX Hoogenboom HRJM, Henderikx MPG;

XX

XX WPI; 2001-626437/72.

XX Novel isolated tumor-associated antigen mucin-1-specific binding member  
 PT for diagnosing and treating cancer, comprises mu-1 binding domain or  
 PT its portion for binding to an epitope of the protein core of mucin-1  
 XX  
 PS Example 2; Page 119; 126pp; English.  
 XX  
 CC The invention relates to an isolated tumour-associated antigen mucin-1  
 CC (MUC-1)-specific binding member comprising an antigen binding domain  
 CC region having an antibody variable light (VL) or heavy (VH) region,  
 CC or a complementarity determining region (CDR) of VL or VH. MUC1-specific  
 CC binding member is useful for diagnosing cancer, preferably adenocarcinoma  
 CC The binding of MUC1-specific binding member to MUC1 is detected by a  
 CC detection method selected from enzyme-linked immunosorbent assay,  
 CC magnetic resonance imaging, scintillation counting, and x-ray film.  
 CC MUC1-specific binding member is useful for treating cancer, preferably  
 CC adenocarcinoma, in an individual, where the cancer is present in tissue  
 CC of the breast, ovary, lung, or bladder of the individual. MUC1-specific  
 CC binding member is useful for diagnosing and imaging MUC1-expressing  
 CC cancer cells and tissues, for purifying or isolating non-glycosylated,  
 CC underglycosylated or cancer-associated forms of MUC1 or MUC1 epitope-  
 CC containing molecules, and for therapeutically or prophylactically  
 CC treating cancer. The present sequence is human PH1 Fab antibody VH  
 CC region FR (framework region)3-CDR3 variant.  
 XX  
 SQ Sequence 14 AA;

Query Match 83.3%; Score 45; DB 22; Length 14;  
 Best Local Similarity 77.8%; Pred. No. 0.24;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 HTGXGVWXP 11  
 ||| ||| |  
 Db 3 htgggvwdp 11

RESULT 12  
 AAE12741  
 ID AAE12741 standard; peptide; 14 AA.  
 XX  
 AC AAE12741;  
 XX  
 DT 04-JAN-2002 (first entry)  
 XX  
 DE Human PH1 Fab antibody VH region FR3-CDR3 variant #8.  
 XX  
 KW Human; tumour-associated antigen mucin-1; MUC-1; adenocarcinoma; VH;  
 KW variable heavy chain region; cancer; breast; ovary; lung; bladder;  
 KW cytostatic; therapy; PH1 antibody; variant.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Region 1..2  
 FT /label= Framework\_region\_3  
 FT Region 3..14  
 FT /label= Complementarity\_determining\_region\_3  
 XX  
 PN WO200175110-A2.  
 XX  
 PD 11-OCT-2001.  
 XX  
 PF 30-MAR-2001; 2001WO-US10589.  
 XX  
 PR 30-MAR-2000; 2000US-0538913.  
 XX  
 PA (DYAX-) DYAX CORP.  
 XX  
 PI Hoogenboom HRJM, Henderikx MPG;  
 XX  
 DR WPI; 2001-626437/72.  
 XX

PT Novel isolated tumor-associated antigen mucin-1-specific binding member  
 PT for diagnosing and treating cancer, comprises mu-1 binding domain or  
 XX its portion for binding to an epitope of the protein core of mucin-1  
 XX  
 PS Example 2; Page 119; 126pp; English.  
 XX  
 CC The invention relates to an isolated tumour-associated antigen mucin-1  
 CC (MUC-1)-specific binding member comprising an antigen binding domain  
 CC region having an antibody variable light (VL) or heavy (VH) region,  
 CC or a complementarity determining region (CDR) of VL or VH. MUC1-specific  
 CC binding member is useful for diagnosing cancer, preferably adenocarcinoma  
 CC The binding of MUC1-specific binding member to MUC1 is detected by a  
 CC detection method selected from enzyme-linked immunosorbent assay,  
 CC magnetic resonance imaging, scintillation counting, and x-ray film.  
 CC MUC1-specific binding member is useful for treating cancer, preferably  
 CC adenocarcinoma, in an individual, where the cancer is present in tissue  
 CC of the breast, ovary, lung, or bladder of the individual. MUC1-specific  
 CC binding member is useful for diagnosing and imaging MUC1-expressing  
 CC cancer cells and tissues, for purifying or isolating non-glycosylated,  
 CC underglycosylated or cancer-associated forms of MUC1 or MUC1 epitope-  
 CC containing molecules, and for therapeutically or prophylactically  
 CC treating cancer. The present sequence is human PH1 Fab antibody VH  
 CC region FR (framework region)3-CDR3 variant.  
 XX  
 SQ Sequence 14 AA;

Query Match 83.3%; Score 45; DB 22; Length 14;  
 Best Local Similarity 77.8%; Pred. No. 0.24;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 HTGXGVWXP 11  
 ||| ||| |  
 Db 3 htgggvwdp 11

RESULT 13  
 AAE12742  
 ID AAE12742 standard; peptide; 14 AA.  
 XX  
 AC AAE12742;  
 XX  
 DT 04-JAN-2002 (first entry)  
 XX  
 DE Human PH1 Fab antibody VH region FR3-CDR3 variant #9.  
 XX  
 KW Human; tumour-associated antigen mucin-1; MUC-1; adenocarcinoma; VH;  
 KW variable heavy chain region; cancer; breast; ovary; lung; bladder;  
 KW cytostatic; therapy; PH1 antibody; variant.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Region 1..2  
 FT /label= Framework\_region\_3  
 FT Region 3..14  
 FT /label= Complementarity\_determining\_region\_3  
 XX  
 PN WO200175110-A2.  
 XX  
 PD 11-OCT-2001.  
 XX  
 PF 30-MAR-2001; 2001WO-US10589.  
 XX  
 PR 30-MAR-2000; 2000US-0538913.  
 XX  
 PA (DYAX-) DYAX CORP.  
 XX  
 PI Hoogenboom HRJM, Henderikx MPG;  
 XX  
 DR WPI; 2001-626437/72.  
 XX  
 PT Novel isolated tumor-associated antigen mucin-1-specific binding member

PT for diagnosing and treating cancer, comprises mucin-1 binding domain or  
 PT its portion for binding to an epitope of the protein core of mucin-1 -  
 XX  
 PS Example 2; Page 119; 126pp; English.  
 XX  
 CC The invention relates to an isolated tumour-associated antigen mucin-1  
 CC (MUC-1)-specific binding member comprising an antigen binding domain  
 CC region having an antibody variable light (VL) or heavy (VH) region,  
 CC or a complementarity determining region (CDR) of VL or VH. MUC1-specific  
 CC binding member is useful for diagnosing cancer, preferably adenocarcinoma  
 CC The binding of MUC1-specific binding member to MUC1 is detected by a  
 CC detection method selected from enzyme-linked immunosorbent assay,  
 CC magnetic resonance imaging, scintillation counting, and X-ray film.  
 CC MUC1-specific binding member is useful for treating cancer, preferably  
 CC adenocarcinoma, in an individual, where the cancer is present in tissue  
 CC of the breast, ovary, lung, or bladder of the individual. MUC1-specific  
 CC binding member is useful for diagnosing and imaging MUC1-expressing  
 CC cancer cells and tissues, for purifying or isolating non-glycosylated,  
 CC underglycosylated or cancer-associated forms of MUC1 or MUC1 epitope-  
 CC containing molecules, and for therapeutically or prophylactically  
 CC treating cancer. The present sequence is human PH1 Fab antibody VH  
 CC region FR (framework region)3-CDR3 variant.  
 XX  
 SQ Sequence 14 AA;

Query Match 83.3%; Score 45; DB 22; Length 14;  
 Best Local Similarity 77.8%; Pred. No. 0.24;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 HTGXGVWXP 11  
 ||| ||| |  
 Db 3 htgggvwdp 11

RESULT 14  
 AAEL2743  
 ID AAEL2743 standard; peptide; 14 AA.  
 XX  
 AC AAEL2743;  
 XX  
 DT 04-JAN-2002 (first entry)  
 XX  
 DE Human PH1 Fab antibody VH region FR3-CDR3 variant #10.  
 XX  
 KW Human; tumour-associated antigen mucin-1; MUC-1; adenocarcinoma; VH;  
 KW variable heavy chain region; cancer; breast; ovary; lung; bladder;  
 KW cytostatic; therapy; PH1 antibody; variant.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Region 1..2  
 FT /label= Framework\_region\_3  
 FT Region 3..14  
 FT /label= Complementarity\_determining\_region\_3  
 XX  
 PN WO200175110-A2.  
 XX  
 PD 11-OCT-2001.  
 XX  
 PF 30-MAR-2001; 2001WO-US10589.  
 XX  
 PR 30-MAR-2000; 2000US-0538913.  
 XX  
 PA (DYAX-) DYAX CORP.  
 XX  
 PI Hoogenboom HRJM, Henderikx MPG;  
 XX  
 DR WPI; 2001-626437/72.  
 XX  
 PT Novel isolated tumor-associated antigen mucin-1-specific binding member  
 PT for diagnosing and treating cancer, comprises mucin-1 binding domain or

PT its portion for binding to an epitope of the protein core of mucin-1 -  
 XX  
 PS Example 2; Page 119; 126pp; English.  
 XX  
 CC The invention relates to an isolated tumour-associated antigen mucin-1  
 CC (MUC-1)-specific binding member comprising an antigen binding domain  
 CC region having an antibody variable light (VL) or heavy (VH) region,  
 CC or a complementarity determining region (CDR) of VL or VH. MUC1-specific  
 CC binding member is useful for diagnosing cancer, preferably adenocarcinoma  
 CC The binding of MUC1-specific binding member to MUC1 is detected by a  
 CC detection method selected from enzyme-linked immunosorbent assay,  
 CC magnetic resonance imaging, scintillation counting, and X-ray film.  
 CC MUC1-specific binding member is useful for treating cancer, preferably  
 CC adenocarcinoma, in an individual, where the cancer is present in tissue  
 CC of the breast, ovary, lung, or bladder of the individual. MUC1-specific  
 CC binding member is useful for diagnosing and imaging MUC1-expressing  
 CC cancer cells and tissues, for purifying or isolating non-glycosylated,  
 CC underglycosylated or cancer-associated forms of MUC1 or MUC1 epitope-  
 CC containing molecules, and for therapeutically or prophylactically  
 CC treating cancer. The present sequence is human PH1 Fab antibody VH  
 CC region FR (framework region)3-CDR3 variant.  
 XX  
 SQ Sequence 14 AA;

Query Match 83.3%; Score 45; DB 22; Length 14;  
 Best Local Similarity 77.8%; Pred. No. 0.24;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 HTGXGVWXP 11  
 ||| ||| |  
 Db 3 htgggvwdp 11

RESULT 15  
 AAEL2744  
 ID AAEL2744 standard; peptide; 14 AA.  
 XX  
 AC AAEL2744;  
 XX  
 DT 04-JAN-2002 (first entry)  
 XX  
 DE Human PH1 Fab antibody VH region FR3-CDR3 variant #11.  
 XX  
 KW Human; tumour-associated antigen mucin-1; MUC-1; adenocarcinoma; VH;  
 KW variable heavy chain region; cancer; breast; ovary; lung; bladder;  
 KW cytostatic; therapy; PH1 antibody; variant.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Region 1..2  
 FT /label= Framework\_region\_3  
 FT Region 3..14  
 FT /label= Complementarity\_determining\_region\_3  
 XX  
 PN WO200175110-A2.  
 XX  
 PD 11-OCT-2001.  
 XX  
 PF 30-MAR-2001; 2001WO-US10589.  
 XX  
 PR 30-MAR-2000; 2000US-0538913.  
 XX  
 PA (DYAX-) DYAX CORP.  
 XX  
 PI Hoogenboom HRJM, Henderikx MPG;  
 XX  
 DR WPI; 2001-626437/72.  
 XX  
 PT Novel isolated tumor-associated antigen mucin-1-specific binding member  
 PT for diagnosing and treating cancer, comprises mucin-1 binding domain or  
 PT its portion for binding to an epitope of the protein core of mucin-1 -



XX Example 2; Page 119; 126pp; English.  
PS  
XX  
CC The invention relates to an isolated tumour-associated antigen mucin-1  
CC (MUC-1)-specific binding member comprising an antigen binding domain  
CC region having an antibody variable light (VL) or heavy (VH) region,  
CC or a complementarity determining region (CDR) of VL or VH. MUC1-specific  
CC binding member is useful for diagnosing cancer, preferably adenocarcinoma  
CC The binding of MUC1-specific binding member to MUC1 is detected by a  
CC detection method selected from enzyme-linked immunosorbent assay,  
CC magnetic resonance imaging, scintillation counting, and X-ray film.  
CC MUC1-specific binding member is useful for treating cancer, preferably  
CC adenocarcinoma, in an individual, where the cancer is present in tissue  
CC of the breast, ovary, lung, or bladder of the individual. MUC1-specific  
CC binding member is useful for diagnosing and imaging MUC1-expressing  
CC cancer cells and tissues, for purifying or isolating non-glycosylated,  
CC underglycosylated or cancer-associated forms of MUC1 or MUC1 epitope-  
CC containing molecules, and for therapeutically or prophylactically  
CC treating cancer. The present sequence is human PH1 Fab antibody VH  
CC region FR (framework region)3-CDR3 variant.  
XX  
SQ Sequence 14 AA;

Query Match 83.3%; Score 45; DB 22; Length 14;  
Best Local Similarity 77.8%; Pred. No. 0.24;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 HTGXGVWXP 11  
||| ||| |  
Db 3 htgggvwdp 11

Search completed: October 4, 2002, 10:21:47  
Job time: 265 sec

Page 10

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 4, 2002, 10:28:44 ; Search time 15.77 Seconds  
(without alignments)  
34.374 Million cell updates/sec

Title: US-09-822-698A-28  
Perfect score: 54  
Sequence: 1 XXHTGXGVWXPXXX 14

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	38	70.4	250	1 ETV3_HUMAN	P41162 homo sapien
2	38	70.4	656	1 RIP_MOUSE	Q08055 mus musculus
3	38	70.4	807	1 YAKI_YEAST	P34680 saccharomyc
4	37	68.5	537	1 ACHI_SCHGR	P24814 schistocerc
5	36	66.7	567	1 ACHI_DROME	P09478 drosophila
6	34	63.0	473	1 ARSE_RAT	P50430 rattus norv
7	34	63.0	762	1 FIBA_RAT	P06399 rattus norv
8	33	61.1	114	1 KDGL_BACSU	P19638 bacillus su
9	33	61.1	413	1 PURK_CORAM	Q44678 corynebacte
10	33	61.1	478	1 OCT2_PIG	Q29013 sus scrofa
11	33	61.1	487	1 ACH6_CAEEL	Q93149 caenorhabdi
12	33	61.1	536	1 GVDI_HALNI	P13043 halobacteri
13	33	61.1	541	1 Y537_SYNY3	P54148 synechocyst
14	33	61.1	2103	1 RRPL_UUK	P33453 uukuniemi v
15	33	61.1	2245	1 MYSJ_DICDI	P54697 dictyosteli
16	32	59.3	456	1 ACHA_BRARE	Q98880 brachydanio
17	32	59.3	456	1 ACHA_CHICK	P09479 gallus gall
18	32	59.3	573	1 ZOID_COMTE	Q06401 comanonas t
19	32	59.3	697	1 CEAD_ECOLI	P17998 escherichia
20	32	59.3	864	1 AGUO_MUCJA	Q92442 mucor javan
21	32	59.3	1033	1 CR2_HUMAN	P20023 homo sapien
22	32	59.3	1279	1 APU_THESA	P36905 t amylopull
23	32	59.3	1475	1 APU_THEFT	P16950 t amylopull
24	32	59.3	1481	1 APU_THEFT	P16950 t amylopull
25	32	59.3	1861	1 APU_THETU	P38536 t amylopull
26	31	57.4	62	1 YCF9_PORPU	P13136 porphyra pu
27	31	57.4	126	1 CRGB_DEIRA	Q9rx00 deinococcus
28	31	57.4	142	1 YAI6_HAEIN	P44095 haemophilus
29	31	57.4	147	1 AFAD_ECOLI	Q47038 escherichia
30	31	57.4	182	1 AHT1_YEAST	P29589 saccharomyc
31	31	57.4	234	1 POLG_FMDVI	P03310 foot-and-mo
32	31	57.4	255	1 COX3_ASCSU	P24879 ascaris suu
33	31	57.4	262	1 LAT_HUMAN	O43561 homo sapien

34	31	57.4	390	1 PE23_HUMAN	P43115 homo sapien
35	31	57.4	493	1 ACH6_TORCA	P02712 torpedo cal
36	31	57.4	494	1 PACB_BPI1	P27753 bacterioph
37	31	57.4	517	1 CRTI_MYXXA	P54979 myxococcus
38	31	57.4	575	1 MSPA_TREMA	Q92413 treponema m
39	31	57.4	576	1 ACH2_DROME	P17644 drosophila
40	31	57.4	682	1 E13B_BACCI	P23903 bacillus ci
41	31	57.4	715	1 ATI2_HSV1F	P08314 herpes simp
42	31	57.4	718	1 ATI2_HSV11	P10230 herpes simp
43	31	57.4	782	1 L100_ADEL12	P36714 human adeno
44	31	57.4	820	1 CHIA_ALTSO	P32823 alteromonas
45	31	57.4	861	1 POLG_FMDVS	P03311 foot-and-mo

## ALIGNMENTS

RESULT 1					
ETV3_HUMAN					
ID	ETV3_HUMAN	STANDARD;	PRT;	250 AA.	
AC	P41162;				
DT	01-FEB-1995 (Rel. 31, Created)				
DT	01-FEB-1995 (Rel. 31, Last sequence update)				
DT	16-OCT-2001 (Rel. 40, Last annotation update)				
DE	ETS-related protein PE-1 (ETS translocation variant 3) (Fragment).				
GN	ETV3 OR PEL.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Muteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominid				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=94292214; PubMed=8020980;				
RA	Klensz M., Hromas R., Raskind W., Bruno E., Hoffner R.,				
RT	"PE-1, a novel ETS oncogene family member, localized to chromosome				
RT	1q21-q23."				
RL	Genomics 20:291-294(1994).				
CC	- SURCELLULAR LOCATION: Nuclear (Potential).				
CC	- SURCELLULAR LOCATION: BELONGS TO THE ETS FAMILY.				
CC	- This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@isb-sib.ch).				
CC	-----				
DR	EMBL; L16464; AAA60949.1; -				
DR	HSSP; Q01543; IFLI.				
DR	MIM; 164873; -				
DR	InterPro; IPR000418; Ets.				
DR	InterPro; IPR002341; HSF_ETS.				
DR	Pfam; PF00178; Ets; 1.				
DR	SMART; SM00413; Ets; 1.				
DR	PROSITE; PS00345; ETS_DOMAIN_1; 1.				
DR	PROSITE; PS00346; ETS_DOMAIN_2; 1.				
DR	PROSITE; PS50061; ETS_DOMAIN_3; 1.				
KW	DNA-binding; Nuclear protein.				
FT	DNA_BIND 1 138 ETS-DOMAIN.				
FT	DNA_BIND 57 138				
SQ	SEQUENCE 250 AA; 28396 MW; F523E0B2789D0C5 CEM:54;				

Query Match 70.4%; Score 38; DB 1; Length 250;  
Best Local Similarity 75.0%; Pred. No. 4.7;  
Matches 6; Conservative 0; Mismatches 2; Gaps 0;

QY 4 TGXGVWXP 11  
DB 223 TGAGVWIP 230

```
RESULT 2
RIP_MOUSE STANDARD; PRT; 656 AA.
AC Q60855;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE serine/threonine protein kinase RIP (EC 2.7.1.-) (Cell death protein
DE RIP) (Receptor interacting protein).
GN RIPK1 OR RIP OR RINP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=C57BL/6 X CBA; TISSUE=Thymus;
RX MEDLINE=95277838; PubMed=7538908;
RA Stanger B.Z., Leder P., Lee T.-H., Kim E., Seed B.;
RT "RIP: a novel protein containing a death domain that interacts with
RT Fas/Apo-1 (CD95) in yeast and causes cell death.";
RL Cell 81:513-523(1995).
CC -!- FUNCTION: INTERACTS WITH THE DEATH DOMAIN OF FAS AND TRADD AND
CC INITIATES APOPTOSIS. IT IS RECRUITED BY TRADD TO TNFR1 IN A TNF-
CC DEPENDENT PROCESS. REQUIRED FOR TNFR1 ACTIVATION OF NF-KAPPA B.
CC -!- TISSUE SPECIFICITY: FOUND AT LOW LEVELS IN ALL TISSUES.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC -!- SIMILARITY: CONTAINS 1 DEATH DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U25995; AAB60487.1; -.
CC DR HSP; P11362; IAGW.
CC DR MGD; MG1:108212; Ripk1.
CC DR InterPro; IPR000488; Death.
CC DR InterPro; IPR000719; Euk_pkinase.
CC DR InterPro; IPR004040; STY_pkinase.
CC DR Pfam; PF00531; death; 1.
CC DR Pfam; PF00069; pkinase; 1.
CC DR SMART; SM00005; DEATH; 1.
CC DR SMART; SM00221; STYKC; 1.
CC DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
CC DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
CC DR PROSITE; PS50017; DEATH_DOMAIN; 1.
CC DR Transferase; Serine/threonine-protein kinase; ATP-binding;
KW Apoptosis.
FT DOMAIN 17 290
FT NP_BIND 23 31
FT BINDING 49 49
FT ACT_SITE 138 138
FT DOMAIN 568 654
FT VARIANT 473 473 T -> I.
SQ SEQUENCE 656 AA; 74854 MW; ABB350B523879933 CRC64;

Query Match 70.48; Score 38; DB 1; Length 656;
Best Local Similarity 75.08; Pred. No. 12;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 TGXGVWXP 11
Db 473 TGTGVWYP 480

RESULT 3
YAK1_YEAST STANDARD; PRT; 807 AA.
AC P14680;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Protein kinase YAK1 (EC 2.7.1.-).
GN YAK1 OR YUL141C OR J0652.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycos.
OX NCBI_TaxID=4932;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=90108683; PubMed=2558053;
RX MEDLINE=90108683; PubMed=2558053;
RA Garrett S., Broach J.;
RT "Loss of Ras activity in Saccharomyces cerevisiae is suppressed by
RT disruptions of a new kinase gene, YAK1, whose product may act
RT downstream of the camp-dependent protein kinase.";
RL Genes Dev. 3:1336-1348(1989).
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN=S288C / FY1679;
RX MEDLINE=96408771; PubMed=8813765;
RA Katsoulou C., Tzernia M., Tavernarakis N., Alexandaki D.;
RT "Sequence analysis of a 40.7 kb segment from the left arm of yeast
RT chromosome X reveals 14 known genes and 13 new open reading frames
RT including homologues of genes clustered on the right arm of
RT chromosome XI.";
RL Yeast 12:787-797(1996).
CC -!- FUNCTION: THIS PROTEIN OCCURS IN REVERTANT RAS/CAMP MUTANTS
CC YEAST AND MAY BE INVOLVED IN CELL-CYCLE REGULATION.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC MNE/DYRK SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X16056; CAA34192.1; -.
CC DR EMBL; X87371; CAA60814.1; -.
CC DR EMBL; Z49417; CAA89437.1; -.
CC DR PIR; A32582; A32582.
CC DR HSP; P24941; ICKP.
CC DR SGD; S0003677; YAK1.
CC DR InterPro; IPR000719; Euk_pkinase.
CC DR InterPro; IPR002290; Ser_thr_pkinase.
CC DR Pfam; PF00069; pkinase; 2.
CC DR SMART; SM00220; STYKC; 1.
CC DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
CC DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
CC DR Transferase; Serine/threonine-protein kinase; ATP-binding;
KW Phosphorylation.
FT DOMAIN 56 85
FT NP_BIND 369 704
FT BINDING 375 383
FT BINDING 398 398
FT ACT_SITE 496 496
FT MOD_RES 127 127
FT MOD_RES 206 206
FT MOD_RES 240 240
FT MOD_RES 295 295
SQ SEQUENCE 807 AA; 91245 MW; E0B7C56FAA35E056 CRC64;

Query Match 70.48; Score 38; DB 1; Length 807;
Best Local Similarity 55.68; Pred. No. 14;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
```

OY 3 HTGSGVWXP 11  
| | | | |  
DB 43 HMGRIWNP 51

## RESULT 4

ACH1\_SCHGR  
ID ACH1\_SCHGR STANDARD; PRT; 557 AA.  
AC P23414;  
DT 01-NOV-1991 (Rel. 20, Created)  
DT 01-NOV-1991 (Rel. 20, Last sequence update)  
DT 01-FEB-1994 (Rel. 28, Last annotation update)  
DE Acetylcholine receptor protein, alpha-L1 chain precursor.  
OS Schistocerca gregaria (Desert locust).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Insecta;  
OC Acridomorpha; Acridoidea; Acrididae; Schistocerca.  
OX NCBI\_TaxID=7010;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91092263; PubMed=1702381;  
RA Marshall J., Buckingham S.D., Shingai R., Goosey M.W.,  
RA Darlison M.G., Sattelle D.B., Barnard E.A.;  
RT "Sequence and functional expression of a single alpha subunit of an  
RT insect nicotinic acetylcholine receptor.";  
RL EMBL J. 9:4391-4398(1990).  
CC -!- FUNCTION: AFTER BINDING ACETYLCHOLINE, THE ACHR RESPONDS BY AN  
CC EXTENSIVE CHANGE IN CONFORMATION THAT AFFECTS ALL SUBUNITS AND  
CC LEADS TO OPENING OF AN ION-CONDUCTING CHANNEL ACROSS THE PLASMA  
CC MEMBRANE.  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
CC -!- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; X55439; CAA39081.1; -  
CC PIR; S12359; S12359.  
CC InterPro; IPR000188; GABAA\_receptor.  
CC InterPro; IPR001175; Neur\_channel.  
CC Pfam; PF02931; Neur\_chan\_LBD; 1.  
CC Pfam; PF02932; Neur\_chan\_memb; 1.  
CC PRINTS; PR00252; NRIONCHANNEL.  
CC PROSITE; PS00236; NEUROTR\_ION\_CHANNEL; 1.  
CC Receptor; Postsynaptic membrane; Ionic channel; Glycoprotein; Signal;  
CC Transmembrane; Multigene family.  
CC SIGNAL 1 23 POTENTIAL.  
CC CHAIN 24 557 ACETYLCHOLINE RECEPTOR PROTEIN, ALPHA-L1  
CC DOMAIN 24 244  
CC TRANSMEM 245 266  
CC TRANSMEM 274 294  
CC TRANSMEM 308 329  
CC DOMAIN 330 500  
CC TRANSMEM 501 523  
CC DISULFID 151 165  
CC DISULFID 224 225  
CC BY SIMILARITY.  
CC ASSOCIATED WITH RECEPTOR ACTIVATION  
CC (BY SIMILARITY).  
CC N-LINKED (GLCNAC. . .) (POTENTIAL).  
CC N-LINKED (GLCNAC. . .) (POTENTIAL).  
CC ALA/SER-RICH.  
CC HIS-RICH.  
CC SEQUENCE 557 AA; 63026 MW; 168389C887DFDF3E CRC64;

Query Match 68.5%; Score 37; DB 1; Length 557;  
Best Local Similarity 66.7%; Pred. No. 15;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 3 HTGSGVWXP 11  
| | | | |  
DB 135 HTGKVVWTP 143

## RESULT 5

ACH1\_DROME  
ID ACH1\_DROME STANDARD; PRT; 567 AA.  
AC P09478;  
DT 01-MAR-1989 (Rel. 10, Created)  
DT 01-MAR-1989 (Rel. 10, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE Acetylcholine receptor protein, alpha-like chain precursor.  
GN ACRB OR ALS OR ACR96AA.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=OREGON-R;  
RX MEDLINE=88283626; PubMed=2840281;  
RA Bossy B., Ballivet M., Spierer P.;  
RT "Conservation of neural nicotinic acetylcholine receptors from  
RT Drosophila to vertebrate central nervous systems.";  
RL EMBL J. 7:611-618(1988).  
CC -!- FUNCTION: AFTER BINDING ACETYLCHOLINE, THE ACHR RESPONDS BY AN  
CC EXTENSIVE CHANGE IN CONFORMATION THAT AFFECTS ALL SUBUNITS AND  
CC LEADS TO OPENING OF AN ION-CONDUCTING CHANNEL ACROSS THE PLASMA  
CC MEMBRANE.  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
CC -!- TISSUE SPECIFICITY: CNS IN EMBRYOS.  
CC -!- DEVELOPMENTAL STAGE: LATE EMBRYONIC, LATE PUPAL, SECOND INSTAR  
CC LARVAE STAGES.  
CC -!- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; X07194; CAA30172.1; -  
CC PIR; S00381; ACFFAL.  
CC Flybase; FBgn0000036; nAChR-alpha-96Aa.  
CC InterPro; IPR000188; GABAA\_receptor.  
CC InterPro; IPR001175; Neur\_channel.  
CC Pfam; PF02931; Neur\_chan\_LBD; 1.  
CC Pfam; PF02932; Neur\_chan\_memb; 1.  
CC PRINTS; PR00252; NRIONCHANNEL.  
CC PROSITE; PS00236; NEUROTR\_ION\_CHANNEL; 1.  
CC Receptor; Postsynaptic membrane; Ionic channel; Glycoprotein; Signal;  
CC Transmembrane; Multigene family; Polymorphism.  
CC SIGNAL 1 21  
CC CHAIN 22 567 ACETYLCHOLINE RECEPTOR PROTEIN, ALPHA-  
CC LIKE CHAIN 1.  
CC EXTRACELLULAR.  
CC DOMAIN 22 240  
CC TRANSMEM 241 264  
CC TRANSMEM 272 290  
CC TRANSMEM 306 325  
CC DOMAIN 326 513  
CC DOMAIN 514 532  
CC DISULFID 149 163  
CC DISULFID 222 223  
CC BY SIMILARITY.  
CC ASSOCIATED WITH RECEPTOR ACTIVATION  
CC (BY SIMILARITY).  
CC N-LINKED (GLCNAC. . .) (POTENTIAL).  
CC N-LINKED (GLCNAC. . .) (POTENTIAL).  
CC H -> Y.  
CC SEQUENCE 567 AA; 64016 MW; 35C84F67E907D4AE CRC64;

Query Match 66.7%; Score 36; DB 1; Length 567;  
Best Local Similarity 66.7%; Pred. NO. 24;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 HTGXGVWXP 11  
||| |  
Db 133 HTGKVWKP 141

```
RESULT 6
FIBR_RAT
ID ARSB_RAT STANDARD; PRT; 473 AA.
AC P50430;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Arylsulfatase B (EC 3.1.6.12) (ASB) (N-acetylgalactosamine-
  4-sulfatase) (G4S) (Fragment).
GN ARSB
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Liver;
RX MEDLINE=96121368; PubMed=8575749;
RA Kunieda T.;
RT "Mucopolysaccharidosis type VI in rats: isolation of cDNAs encoding
  arylsulfatase B, chromosomal localization of the gene, and
  identification of the mutation.";
RL Genomics 29:582-587(1995).
CC -1- CATALYTIC ACTIVITY: Hydrolysis of the 4-sulfate groups of the N-
  acetyl-D-galactosamine 4-sulfate units of chondroitin sulfate and
  dermatan sulfate.
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Lysosomal.
CC -1- DISEASE: DEFECTS IN ARSB ARE THE CAUSE OF MAROTEAUX-LAMY
  SYNDROME; ALSO KNOWN AS MUCOPOLYSACCHARIDOSIS TYPE VI (MPS-VI).
CC -1- SIMILARITY: BELONGS TO THE SULFATASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation
  at the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D49434; BAA08412.1; -.
DR HSP; P15848; IFSU.
DR InterPro; IPR000917; Sulfatase.
DR Pfam; PF00884; Sulfatase; 1.
DR PROSITE; PS00523; SULFATASE_1; 1.
DR PROSITE; PS00149; SULFATASE_2; FALSE_NEG.
KW Hydrolase; Glycoprotein; Lysosome; Mucopolysaccharidosis.
FT MOD_RES 1 1
FT MOD_RES 31 31 2-AMINO-3-OXOPROPIONIC ACID (BY
  SIMILARITY).
FT ACT_SITE 87 87
FT DISULFID 57 461
FT DISULFID 61 95
FT DISULFID 121 132
FT DISULFID 345 387
FT CARBOHYD 128 128 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 219 219 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 231 231 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 355 355 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 366 366 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 398 398 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT SEQUENCE 473 AA; 53320 MW; 4E114C923A24AF8F CRC64;
```

Query Match 63.0%; Score 34; DB 1; Length 473;  
Best Local Similarity 71.4%; Pred. NO. 46;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 GXGVWXP 11  
|||||  
Db 465 GTGVWSP 471

```
RESULT 7
FIBR_RAT
ID FIBR_RAT STANDARD; PRT; 782 AA.
AC P06399;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Fibrinogen alpha/alpha-E chain precursor [Contains: Fibrinopeptide A].
GN FGA.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC Fu Y., Cao Y., Hertzberg K., Grieninger G.;
  Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
RL [2]
RX MEDLINE=86011580; PubMed=4046033;
RA Crabtree G.R.; Comeau C.M.; Fowlkes D.M.; Fornace A.J. Jr.;
  Malley J.D.; Kant J.A.;
RT "Evolution and structure of the fibrinogen genes. Random insertion of
  introns or selective loss";
RL J. Mol. Biol. 185:1-19(1985).
RN [3]
RP SEQUENCE OF 20-36.
RA Blomback B., Blomback M., Grondahl N.J.;
RT "Studies on fibrinopeptides from mammals.";
RL Acta Chem. Scand. 19:1789-1791(1965).
RN [4]
RP SEQUENCE OF 458-550 FROM N.A. (ALPHA FORM).
RC STRAIN=WISTAR; TISSUE=Liver;
RX MEDLINE=87134033; PubMed=3817019;
RA Sobczak J., Lotti A.-M., Taroux P., Duguet M.;
RT "Molecular cloning of mRNA sequences transiently induced during rat
  liver regeneration.";
RL Exp. Cell Res. 169:47-56(1987).
CC -1- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT
  POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET
  AGGREGATION.
CC -1- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
  (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
  THE AMINO ENDS OF ALL CHAINS ARE CONTAINED IN THE CENTRAL NODULE.
  DIVERGING FROM THIS NODULE ARE 2 THREE-CHAIN COILED COILS, WHICH
  CONNECT THE CENTRAL NODULE TO THE DISTAL NODULES CONTAINING THE
  DISTAL DOMAINS. EXTENDING FAR PERIPHERALLY ARE THE LONG CARBOXYL
  ENDS OF THE ALPHA CHAINS.
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; ALPHA-E AND ALPHA; ARE
  PRODUCED BY ALTERNATIVE SPLICING. THE ALPHA FORM IS THE
  PREDOMINANT FORM.
CC -1- PTM: THE ALPHA CHAIN IS NOT GLYCOSYLATED.
CC -1- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY
  THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA
  CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES
  RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT. THE SOFT CLOT IS
  CONVERTED INTO THE HARD CLOT BY FACTOR XIIIa WHICH CATALYZES THE
  EPSTON-(GAMMA-GLUTAMYL)LYSINE CROSS-LINKING BETWEEN GAMMA CHAINS
  (STRONGER) AND BETWEEN ALPHA CHAINS (WEAKER) OF DIFFERENT
  MONOMERS.
CC -1- SIMILARITY: CONTAINS 1 FIBRINOGEN C-TERMINAL DOMAIN.
```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 DR EMBL; X86561; CAA60264.1; -;  
 DR EMBL; X86561; CAA60263.1; -;  
 DR EMBL; M35601; AAA41158.1; -;  
 DR PIR; A03119; FGTA.  
 DR HSP; P02671; LFZD;  
 DR InterPro; IPR002181; Fibrinogen\_C.  
 DR Pfam; PF00147; fibrinogen\_C; 1.  
 DR SMART; SM00186; FBG; 1.  
 DR PROSITE; PS00514; FIBRIN\_AG\_C\_DOMAIN; 1.  
 KW Blood coagulation; Plasma; Platelet; signal; Alternative splicing.  
 FT SIGNAL 1 19 FIBRINOPEPTIDE A.  
 FT PEPTIDE 20 36 FIBRINOPEPTIDE A.  
 FT CHAIN 20 782 FIBRINOGEN ALPHA-E CHAIN.  
 FT SITE 36 37 CLEAVAGE (BY THROMBIN; RELEASE  
 FT FIBRINOPEPTIDE A).  
 FT DISULFID 48 48 INTERCHAIN (WITH C-48') (BY SIMILARITY).  
 FT DISULFID 56 56 INTERCHAIN (WITH BETA) (BY SIMILARITY).  
 FT DISULFID 65 65 INTERCHAIN (WITH C-49 IN GAMMA  
 FT (BY SIMILARITY).  
 FT DISULFID 69 69 INTERCHAIN (WITH BETA) (BY SIMILARITY).  
 FT DISULFID 181 181 INTERCHAIN (WITH C-165 IN GAMMA  
 FT (BY SIMILARITY).  
 FT DISULFID 185 185 INTERCHAIN (WITH BETA) (BY SIMILARITY).  
 FT DISULFID 404 434 BY SIMILARITY.  
 FT VARSPLIC 547 550 DCDD -> GIHA (IN ISOFORM ALPHA).  
 FT VARSPLIC 551 782 MISSING (IN ISOFORM ALPHA).  
 FT CONFLICT 30 34 EAGD -> DEAG (IN REF. 3).  
 FT CONFLICT 140 140 Q -> E (IN REF. 2).  
 FT CONFLICT 212 212 D -> E (IN REF. 2).  
 FT CONFLICT 270 276 ASRGDLF -> LREEIYQ (IN REF. 2).  
 FT CONFLICT 473 473 S -> K (IN REF. 4).  
 SQ SEQUENCE 782 AA; 86685 MW; 744834DAE76D34C2 CRC64;

Query Match 63.08; Score 34; DB 1; Length 782;  
 Best Local Similarity 55.68; Pred. No. 74;  
 Matches 5; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 HTGXGWWXP 11  
 | | | | |  
 Db 300 HGGSGYWRP 308

RESULT 8  
 KDGL\_BACSU  
 ID KDGL\_BACSU STANDARD; PRT; 114 AA.  
 AC P19638;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Diacylglycerol kinase (EC 2.7.1.107) (DAGK) (diglyceride kinase)  
 DE (DAGK).  
 GN DGK OR DGK.  
 OS Bacillus subtilis.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
 OC Bacillus/Staphylococcus group; Bacillus.  
 OX NCBI\_TaxID=1423;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ED40;  
 RA Kim K., Hwang S., Suh J., Song B.-H., Hong S., Kim J.;  
 RT "Nucleotide sequence upstream of the cdd locus in Bacillus subtilis.";  
 RL J. Microbiol. Biotechnol. 5:0-0(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.

RC STRAIN=168 / JH642;  
 RA Kobayashi Y., Mizuno M., Masuda S., Takamaru K., Hirono S.,  
 RA Sato T., Takeuchi M.;  
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE OF 3-114 FROM N.A.  
 RC STRAIN=168;  
 RA MEDLINE=89313687; PubMed=2526291;  
 RX Song B.-H., Neuhard J.;  
 RT "Chromosomal location, cloning and nucleotide sequence of the  
 RT Bacillus subtilis cdd gene encoding cytidine/deoxycytidine  
 RT deaminase.";  
 RL Mol. Gen. Genet. 216:462-468(1989).  
 RN [4]  
 RP IDENTIFICATION.  
 RX MEDLINE=94350829; PubMed=8071224;  
 RA Smith R.L., O'Toole J.F., Maquire M.E., Sanders G.R. II;  
 RT "Membrane topology of Escherichia coli diacylglycerol kinase.";  
 RL J. Bacteriol. 176:5459-5465(1994).  
 CC -1- FUNCTION: RECYCLING OF DIACYLGLYCEROL PRODUCED DURING THE  
 CC TURNOVER OF MEMBRANE PHOSPHOLIPID (BY SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: ATP + 1,2-diacylglycerol = ADP + 1,2-  
 CC diacylglycerol 3-phosphate.  
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL DIACYLGLYCEROL KINASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 DR EMBL; U29177; AAA70044.1; ALT\_INIT.  
 DR EMBL; D84432; BRA12480.1; -;  
 DR EMBL; X17430; CAB57857.1; -;  
 DR EMBL; Z99116; CAB14460.1; -;  
 DR EMBL; Z99117; CAB14473.1; -;  
 DR PIR; PE0002; PE0002.  
 DR Subtilist; BG11040; dgka.  
 DR InterPro; IPR000829; DAGK\_prokar.  
 DR Pfam; PF01219; DAGK\_prokar; 1.  
 DR PRODOM; PD010722; DAGK\_prokar; 1.  
 DR PROSITE; PS01069; DAGK\_PROKAR; 1.  
 KW Phospholipid biosynthesis; Transferase; Kinase; Transmembrane;  
 KW Complete proteome.  
 FT DOMAIN 1 26 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 27 42 POTENTIAL.  
 FT TRANSMEM 46 62 POTENTIAL.  
 FT DOMAIN 63 89 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 90 114 POTENTIAL.  
 FT CONFLICT 103 103 C -> S (IN REF. 1).  
 SQ SEQUENCE 114 AA; 12544 MW; DA6269C79834EBC CR 64;  
 Query Match 61.1%; Score 33; DB 1; Length 114;  
 Best Local Similarity 57.1%; Pred. No. 18;  
 Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 HTGXGWW 9  
 | | | | |  
 Db 10 HAGRGW 16

RESULT 9  
 PURK\_CORAM  
 ID PURK\_CORAM STANDARD; PRT; 413 AA.  
 AC Q44678;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE Phosphoribosylaminoimidazole carboxylase ATPase subunit (EC 4.1.1.21)  
 DE (AIR carboxylase) (AIRC).

GN PURK.  
OS Corynebacterium ammoniagenes (Brevibacterium ammoniagenes).  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteriidae;  
OC Actinomycetales; Corynebacterineae; Corynebacteriaceae;  
OC Corynebacterium.  
OX NCBI\_TaxID=1697;  
[1]  
RN RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 6872; PubMed=9989996;  
RX MEDLINE=9623688; Lee S.Y., Lee D.S.;  
RA Chung S.O., Lee J.H., Lee S.Y., Lee D.S.;  
RT "Genomic organization of purk and pure in Brevibacterium ammoniagenes  
ATCC 6872: pure locus provides a clue for genomic evolution.";  
RL FEMS Microbiol. Lett. 137:265-268(1996).  
CC -!- FUNCTION: POSSESSES AN APASE ACTIVITY THAT IS DEPENDENT ON THE  
CC PRESENCE OF AIR (AMINOIMIDAZOLE RIBONUCLEOTIDE). THE ASSOCIATION  
CC OF PURK AND PURE PRODUCES AN ENZYME COMPLEX CAPABLE OF CONVERTING  
CC AIR TO CAIR EFFICIENTLY UNDER PHYSIOLOGICAL CONDITION  
CC (BY SIMILARITY).  
CC -!- SIMILARITY: BELONGS TO THE PURK / PURT FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
-----  
DR EMBL; X91189; CAA62598.1; -.  
DR HSP; P09029; 1B6R.  
DR InterPro; IPR003135; ATP-grasp.  
DR Pfam; PF02222; ATP-grasp; 1.  
KW Purine biosynthesis; Lyase; Decarboxylase.  
SQ SEQUENCE 413 AA; 44110 MW; F738B230DF5D119A CRC64;  
  
Query Match 61.1%; Score 33; DB 1; Length 413;  
Best Local Similarity 53.6%; Pred. No. 61;  
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
  
QY 3 HTXGVWXP 11  
DB 168 YDCKGVWFP 176  
  
RESULT 10  
OCT2\_PIG  
ID OCT2\_PIG STANDARD; PRT; 478 AA.  
AC Q29013; Q29089; PRT; 478 AA.  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE Octamer-binding transcription factor 2 (OTF-2) (Lymphoid-restricted  
DE immunoglobulin octamer binding protein NF-A2) (OCT-2 factor).  
DE POU2F2 OR OTF2 OR OCT2.  
OS Sus scrofa (Pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
OX NCBI\_TaxID=9823;  
[1]  
RN RP SEQUENCE FROM N.A.  
RC TISSUE=Spleen;  
RX MEDLINE=95030552; PubMed=7943947;  
RA Tuggle C.K., Helm J., Rothschild M.F.;  
RT "Cloning, sequencing and restriction fragment length polymorphism  
RT analysis of a porcine cDNA for OCT2.";  
RL Anim. Genet. 25:141-145(1994).

[2]  
RN RP SEQUENCE OF 210-350 FROM N.A.  
RC STRAIN=DUFROC;  
RX MEDLINE=94095558; PubMed=8270542;  
RA Tuggle C.K.;  
RT "Cloning and sequence analysis of the swine Oct-2 POU-domain genomic  
RT region.";  
RL J. Anim. Sci. 71:3172-3172(1993).  
CC -!- FUNCTION: THIS PROTEIN IS A TRANSCRIPTION FACTOR THAT SPECIFICALLY  
CC BINDS TO THE OCTAMER MOTIF (ATTGTCAT) AND PLAYS AN IMPORTANT  
CC ROLE IN REGULATING TRANSCRIPTION IN A NUMBER OF TISSUES IN  
CC ADDITION TO ACTIVATING IMMUNOGLOBULIN GENE EXPRESSION (BY  
CC SIMILARITY).  
CC -!- SUBCELLULAR LOCATION: Nuclear.  
CC -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN B-CELLS (BY  
CC SIMILARITY).  
CC -!- SIMILARITY: STRONG TO OTHER "POU" TRANSCRIPTION FACTORS. BELONGS  
CC TO CLASS-2 POU.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
-----  
DR EMBL; U00794; AAA80148.1; -.  
DR EMBL; L03842; AAA74657.1; -.  
DR HSP; P09086; 1HDP.  
DR InterPro; IPR001356; Homeobox.  
DR InterPro; IPR000972; Octamer-bind\_tf.  
DR InterPro; IPR000327; POU.  
DR Pfam; PF00046; homeobox; 1.  
DR Pfam; PF00157; pou; 1.  
DR PRINTS; PR00029; OCTAMER.  
DR PROSITE; PS00028; POUDOMAIN.  
DR PRODOM; PD000583; POU; 1.  
DR SMART; SM00389; Hox; 1.  
DR SMART; SM00352; POU; 1.  
DR PROSITE; PS00027; HOMEBOX\_1; 1.  
DR PROSITE; PS00071; HOMEBOX\_2; 1.  
DR PROSITE; PS00035; POU\_1; 1.  
DR PROSITE; PS00465; POU\_2; 1.  
KW Homeobox; DNA-binding; Transcription regulation; Activator;  
KW Nuclear protein.  
FT DOMAIN 199 269 POU.  
FT DNA\_BIND 297 356 HOMEBOX.  
FT DOMAIN 389 410 LEUCINE-ZIPPER.  
FT DOMAIN 77 80 POLY-PRO.  
FT DOMAIN 417 424 GLY-RICH.  
FT DOMAIN 437 440 POLY-PRO.  
FT CONFLICT 212 212 Q -> V (IN REF. 2).  
FT CONFLICT 215 215 I -> M (IN REF. 2).  
FT CONFLICT 222 222 G -> V (IN REF. 2).  
FT CONFLICT 233 233 G -> A (IN REF. 2).  
FT CONFLICT 238 238 Q -> K (IN REF. 2).  
FT CONFLICT 285 285 R -> S (IN REF. 2).  
SQ SEQUENCE 478 AA; 51098 MW; FD196758B603B718 CRC64;  
  
Query Match 61.1%; Score 33; DB 1; Length 478;  
Best Local Similarity 62.5%; Pred. No. 70;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
  
QY 4 TGXGVWXP 11  
DB 466 TGPGLWNP 473  
  
RESULT 11  
ACH6\_CAEEL  
ID ACH6\_CAEEL STANDARD; PRT; 487 AA.





FT NP\_BIND 39 46 ATP (POTENTIAL).  
SQ SEQUENCE 536 AA; 59342 MW; 3FC4E4F8F129F9CD CRC64;

Query Match 61.1%; Score 33; DB 1; Length 536;  
Best Local Similarity 62.5%; Pred. No. 78;  
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 TCGXGVWP 11  
||| |  
Db 247 TCGTWDTP 254

RESULT 13  
Y537\_SYNY3  
ID Y537\_SYNY3 STANDARD; PRT; 541 AA.  
AC P54148;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Putative ammonium transporter sl10537.  
GN SL10537.

OS Synechocystis sp. (strain PCC 6803).  
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.  
OX NCBI\_TaxID=11148;  
RN [1]

RP SEQUENCE FROM N.A.  
RX MEDLINE=96127529; PubMed=8590279;  
RA Kaneko T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,  
RA Sugita M., Tabata S.;  
RT "Sequence analysis of the genome of the unicellular cyanobacterium  
Synechocystis sp. strain PCC6803. I. Sequence features in the 1 Mb  
region from map positions 64% to 92% of the genome.";  
RL DNA Res. 2:153-166(1995).

CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).  
CC -!- SIMILARITY: BELONGS TO THE AMT1/MEP/NRGA FAMILY OF AMMONIUM  
TRANSPORTERS (TC 2.49).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; D64006; BAA10864.1; -  
DR InterPro; IPR001905; Ammonium\_transp.  
DR InterPro; IPR003660; HAMP.  
DR Pfam; PF00909; Ammonium\_transp; 1.  
DR SMART; SM00304; HAMP; 1.  
DR PROSITE; PS01219; AMMONIUM\_TRANS; 1.  
KW Hypothetical protein; Transport; Transmembrane; Complete proteome.  
FT TRANSMEM 6 26 POTENTIAL.  
FT TRANSMEM 44 64 POTENTIAL.  
FT TRANSMEM 86 106 POTENTIAL.  
FT TRANSMEM 117 137 POTENTIAL.  
FT TRANSMEM 161 181 POTENTIAL.  
FT TRANSMEM 203 223 POTENTIAL.  
FT TRANSMEM 235 255 POTENTIAL.  
FT TRANSMEM 260 280 POTENTIAL.  
FT TRANSMEM 283 303 POTENTIAL.  
FT TRANSMEM 316 336 POTENTIAL.  
FT TRANSMEM 356 376 POTENTIAL.  
SQ SEQUENCE 541 AA; 58129 MW; A7A0A5CB2FFD2475 CRC64;

Query Match 61.1%; Score 33; DB 1; Length 541;  
Best Local Similarity 71.4%; Pred. No. 79;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 HTGXGVW 9  
||| |  
.

Db 322 HGGAGVW 328

RESULT 14  
RRPL\_UUK  
ID RRPL\_UUK STANDARD; PRT; 2103 AA.  
AC P33453;  
DT 01-FEB-1994 (Rel. 28, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE RNA-directed RNA polymerase (EC 2.7.7.48) (L protein).  
GN L.

OS Dengue virus (Uuk).  
OC Viruses; ssRNA negative-strand viruses; Bunyaviridae; Phlebovirus.  
OX NCBI\_TaxID=11591;  
RN [1]  
RP SEQUENCE FROM N.A.

RC STRAIN=S23;  
RX MEDLINE=92333259; PubMed=1629699;  
RA Elliott R.M., Dunn E., Simons J.F., Pettersson R.;  
RT "Nucleotide sequence and coding strategy of the Kunjimi virus L RNA  
segment.";  
RL J. Gen. Virol. 73:1745-1752(1992).  
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +  
(RNA)(N).  
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; J10759; BAA01590.1; -  
DR PIR; J01621; J01621  
KW Transferase; Nucleotidyltransferase; RNA-directed RNA polymerase.  
SQ SEQUENCE 2103 AA; 241039 MW; B0EA708451B0B5ED CRC64;

Query Match 61.1%; Score 33; DB 1; Length 2103;  
Best Local Similarity 71.4%; Pred. No. 2.9e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 HTGXGVW 9

Db 1729 YTGQGVW 1735

RESULT 15  
MYSJ\_DICDI  
ID MYSJ\_DICDI STANDARD; PRT; 2245 AA.  
AC P54697;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Myosin IJ heavy chain.  
GN MYOJ

OS Dictyostelium discoideum (Slime mold).  
OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.  
OX NCBI\_TaxID=44689;  
RN [1]  
RP SEQUENCE FROM N.A.

RC STRAIN=AX3;  
RX MEDLINE=96215148; PubMed=8636147;  
RA Hammer J.A. III, Jung G.;

RT "The sequence of the dictyostelium myo J heavy chain gene predicts a  
RT novel, dimeric, unconventional myosin with a heavy chain molecular  
RT mass of 258 kDa.";  
RL J. Biol. Chem. 271:7120-7127(1996).  
RN [2]

RP SEQUENCE OF 1-1021 FROM N.A.  
RX MEDLINE=97039016; PubMed=8884597;

Search completed: October 4, 2002, 10:28:45  
Job time: 387 sec

```
RA Peterson M.D., Urioste A.S., Titus M.A.;  
RT "Dictyostelium discoideum myoJ: a member of a broadly defined myosin  
RL v class or a class xi unconventional myosin?";  
RN J. Muscle Res. Cell Motil. 17:411-424(1996).  
[3]  
RX SEQUENCE OF 182-298 FROM N.A.  
RY MEDLINE-95023928; PubMed-7937787;  
RA Titus M.A., Kuspa A., Loomis W.F.;  
RT "Discovery of myosin genes by physical mapping in Dictyostelium.";  
RL Proc. Natl. Acad. Sci. U.S.A. 91:9446-9450(1994).  
CC -!- FUNCTION: MYOSIN IS A PROTEIN THAT BINDS TO ACTIN AND HAS ATPASE  
CC ACTIVITY THAT IS ACTIVATED BY ACTIN.  
CC -!- SUBUNIT: HOMODIMER.  
CC -!- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.  
CC -!- SIMILARITY: CONTAINS 3 IQ DOMAINS.  
CC -!- SIMILARITY: CONTAINS 1 DILUTE DOMAIN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC -----  
DR EMBL: U42409; AAA85186.1; -;  
DR EMBL: L35322; AAA79858.1; -;  
DR HSSP: P08799; LMND.  
DR DictyDb: DD01095; myoJ.  
DR InterPro: IPR002710; DIL.  
DR InterPro: IPR000048; IQ.  
DR InterPro: IPR004009; Myosin_N.  
DR InterPro: IPR001609; myosin_head.  
DR Pfam: PF01843; DIL; 1.  
DR Pfam: PF00612; IQ; 6.  
DR Pfam: PF00063; myosin_head; 2.  
DR Pfam: PF02736; Myosin_N; 1.  
DR PRINTS: PR00193; MYOSINHEAVY.  
DR ProDom: PD000335; myosin_head; 1.  
DR ProDom: PD003376; DIL; 1.  
DR SMART: SM00015; IQ; 3.  
DR SMART: SM00242; MYSC; 1.  
DR PROSITE: PS50096; IQ; 3.  
KW Myosin: ATP-binding; Actin-binding; Calmodulin-binding; Repeat;  
KW Coiled coil.  
FT DOMAIN 1 809 MYOSIN HEAD-LIKE.  
FT DOMAIN 824 851 IQ 1.  
FT DOMAIN 872 901 IQ 2.  
FT DOMAIN 943 972 IQ 3.  
FT DOMAIN 973 1812 COILED COIL (POTENTIAL).  
FT DOMAIN 1813 2245 TAIL.  
FT NP_BIND 174 181 ATP (POTENTIAL).  
FT DOMAIN 669 749 ACTIN-BINDING.  
FT CONFLICT 191 191 L -> F (IN REF. 2).  
FT CONFLICT 284 284 A -> T (IN REF. 2).  
FT CONFLICT 291 291 G -> R (IN REF. 2).  
FT CONFLICT 332 347 NKSGCFEIEGVSDDEH -> IEWPELKVYRMKS (IN  
FT REF. 2).  
FT CONFLICT 550 550 N -> K (IN REF. 2).  
FT CONFLICT 865 866 HH -> QQ (IN REF. 2).  
SQ SEQUENCE 2245 AA; 258478 MW; 615E5EFID1AB45BE CRC64;
```

Query Match 61.1%; Score 33; DB 1; Length 2245;  
Best Local Similarity 71.4%; Pred. No. 3e+02;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 5 GXGWXP 11  
| | | | |  
Db 27 GAGWIP 33

4835

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 4, 2002, 10:17:22 ; Search time 51.14 Seconds  
(without alignments)  
262.807 Million cell updates/sec

Title: US-09-822-698A-3  
Perfect score: 644  
Sequence: 1 QVQLVQSGGGLVQPGGSLRL.....GGVWMDPIDYWGQGLTVTVSS 121

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_032802.\*  
1: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	644	100.0	121	AAEL2706	Human PH1 Fab anti
2	644	100.0	381	AAEL2707	Human b1vPH1-IL-2
3	644	100.0	451	AAEL2715	Human recombinant
4	548	85.1	121	AAE07029	Human heavy chain
5	543.5	84.4	120	AAW27553	Human Ab heavy cha
6	543.5	84.4	281	AAW27560	Consensus single c
7	541.5	84.1	116	AAE07017	Human heavy chain
8	537.5	83.5	131	AAW13520	Anti-melanoma anti
9	537	83.4	117	AAEL2061	Human anti-tissue
10	537	83.4	313	AAU14320	Human novel protei
11	536.5	83.3	240	AAV02472	A single chain ant

12	535	83.1	117	18	AAW13532	Anti-melanoma anti
13	533	82.8	119	22	AAU02512	Anti-adipocyte mon
14	533	82.8	246	21	AAV58235	Internalising anti
15	532	82.6	227	22	AAW75051	TR0005 Humab Happa
16	531.5	82.5	120	22	AAE07022	Human heavy chain
17	529	82.1	240	21	AAV15124	Anti-human CTLA-4
18	529	82.1	240	21	AAV15125	Anti-murine CTLA-4
19	527.5	81.9	122	20	AAW78432	Antibody heavy cha
20	527	81.8	245	22	AAW67620	Human leukocyte an
21	525	81.5	120	15	AAW54803	SP4-reactive VH re
22	524.5	81.4	116	18	AAW13529	Anti-melanoma anti
23	524	81.4	120	15	AAW54802	SP4-reactive VH re
24	524	81.4	120	15	AAW54804	SP4-reactive VH re
25	524	81.4	120	15	AAW54805	SP4-reactive VH re
26	524	81.4	120	15	AAW54806	SP4-reactive VH re
27	524	81.4	123	22	AAE07019	Human heavy chain
28	523.5	81.3	293	22	AAW65715	Amino acid sequenc
29	523	81.2	113	19	AAW70621	Human consensus fr
30	523	81.2	113	21	AAW82347	Human consensus se
31	523	81.2	119	22	AAW62088	Human Vh consensus
32	523	81.2	119	22	AAW60401	Consensus human he
33	523	81.2	119	22	AAW61586	Human variable hea
34	521.5	81.0	124	22	AAW72879	Human anti-HER2/ne
35	521	80.9	125	22	AAE07013	Human heavy chain
36	520.5	80.8	116	22	AAU02613	Anti-adipocyte mon
37	520.5	80.8	140	18	AAW13524	Anti-melanoma anti
38	520.5	80.7	240	22	AAW60404	Human JF anti-Idio
39	520	80.7	125	15	AAW54784	SP4-reactive IgM h
40	519.5	80.7	118	22	AAW65354	Anti-IL-18 antibod
41	519.5	80.7	120	22	AAU02508	Anti-adipocyte mon
42	519	80.6	125	22	AAE07021	Human heavy chain
43	519	80.6	127	22	AAE07024	Human heavy chain
44	518.5	80.5	128	22	AAE07014	Human heavy chain
45	518	80.4	121	19	AAW47180	Variable heavy (VH)

## ALIGNMENTS

RESULT 1  
AAEL2706  
ID AAEL2706 standard; Protein; 121 AA.  
XX AC AAEL2706;  
XX DT 04-JAN-2002 (first entry)  
XX DE Human PH1 Fab antibody variable heavy chain region (VH).  
XX KW Human: tumour-associated antigen mucin-1; MUC-1; adenocarcinoma; VH;  
variable heavy chain region; cancer; breast; ovary; lung; bladder;  
cytostatic; therapy; PH1 antibody.  
XX OS Homo sapiens.  
XX FH Key Location/Qualifiers  
FH Region 31..35 /label= CDR1  
FH FT /note= "Complementarity determining region 1"  
FH FT 50..66 /label= CDR2  
FH FT /note= "Complementarity determining region 2"  
FH FT 97..99 /label= FR3  
FH FT /note= "Framework region 3"  
FH FT 97..112 /note= "VH domain"  
FH FT 99..110 /label= CDR3  
FH FT /note= "Complementarity determining region 3"  
FH FT 111..112 /label= FR4  
FH FT /note= "Framework region 4"

```

XX PN WO200175110-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US10589.
XX PR 30-MAR-2000; 2000US-0538913.
XX PA (DYAX-) DYAX CORP.
XX PI Hoogenboom HRJM, Henderikx MPG;
XX DR WPI; 2001-626437/72.
XX DR N-PSDB; AAD20731.
XX
XX PT Novel isolated tumor-associated antigen mucin-1-specific binding member
XX PT for diagnosing and treating cancer, comprises mucin-1 binding domain or
XX PT its portion for binding to an epitope of the protein core of mucin-1
XX
XX PS Claim 2; Page 94-95; 126pp; English.
XX
XX CC The invention relates to an isolated tumor-associated antigen mucin-1
XX CC (MUC-1)-specific binding member comprising an antigen binding domain
XX CC region having an antibody variable light (VL) or heavy (VH) region,
XX CC or a complementarity determining region (CDR) of VL or VH. MUC1-specific
XX CC binding member is useful for diagnosing cancer, preferably adenocarcinoma
XX CC The binding of MUC1-specific binding member to MUC1 is detected by a
XX CC detection method selected from enzyme-linked immunosorbent assay,
XX CC magnetic resonance imaging, scintillation counting, and x-ray film.
XX CC MUC1-specific binding member is useful for treating cancer, preferably
XX CC adenocarcinoma, in an individual, where the cancer is present in tissue
XX CC of the breast, ovary, lung, or bladder of the individual. MUC1-specific
XX CC binding member is useful for diagnosing and imaging MUC1-expressing
XX CC cancer cells and tissues, for purifying or isolating non-glycosylated,
XX CC underglycosylated or cancer-associated forms of MUC1 or MUC1 epitope-
XX CC containing molecules, and for therapeutically or prophylactically
XX CC treating cancer. The present sequence is anti-MUC1 human PH1 Fab
XX CC antibody VH region.
XX
XX SQ Sequence 121 AA;

Query Match 100.0%; Score 644; DB 22; Length 121;
Best Local Similarity 100.0%; Pred. No. 2.7e-50;
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QVOLVSGGLVQPGSLRLSCAASGFTFRSNAMGWVRQAPGKLEWVSGISGGSTYY 60
DB 1 qvqlvsgggglvpggslrlscaasgftfrsnamgwvrqapkglewsgisggstyy 60
QY 61 ADSVKGRFTISRDNKNTLYLQMSLRAEDTAVYYCAKHTGGGVWDPIDYWGQGLTVTVS 120
DB 61 adsvkgrftisrdsnknltlylqmslraedtavyycahkhtgggvwdpidywgqglvtvts 120
QY 121 s 121
DB 121 s 121

RESULT 2
AAE12707
ID AAE12707 standard; Protein: 381 AA.
XX
XX AC AAE12707;
XX
XX DT 04-JAN-2002 (first entry)
XX
XX DE Human b1vph1-IL-2 immunocytokine protein.
XX
XX KW Human; tumour-associated antigen mucin-1; MUC-1; adenocarcinoma; cancer;
XX KW breast; ovary; lung; bladder; cytostatic; therapy; immunocytokine.
XX

```

```

OS Homo sapiens.
XX PN WO200175110-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US10589.
XX PR 30-MAR-2000; 2000US-0538913.
XX PA (DYAX-) DYAX CORP.
XX PI Hoogenboom HRJM, Henderikx MPG;
XX DR WPI; 2001-626437/72.
XX DR N-PSDB; AAD20732.
XX
XX PT Novel isolated tumor-associated antigen mucin-1-specific binding member
XX PT for diagnosing and treating cancer, comprises mucin-1 binding domain or
XX PT its portion for binding to an epitope of the protein core of mucin-1
XX
XX PS Claim 9; Page 95-97; 126pp; English.
XX
XX CC The invention relates to an isolated tumor-associated antigen mucin-1
XX CC (MUC-1)-specific binding member comprising an antigen binding domain
XX CC region having an antibody variable light (VL) or heavy (VH) region,
XX CC or a complementarity determining region (CDR) of VL or VH. MUC1-specific
XX CC binding member is useful for diagnosing cancer, preferably adenocarcinoma
XX CC The binding of MUC1-specific binding member to MUC1 is detected by a
XX CC detection method selected from enzyme-linked immunosorbent assay,
XX CC magnetic resonance imaging, scintillation counting, and x-ray film.
XX CC MUC1-specific binding member is useful for treating cancer, preferably
XX CC adenocarcinoma, in an individual, where the cancer is present in tissue
XX CC of the breast, ovary, lung, or bladder of the individual. MUC1-specific
XX CC binding member is useful for diagnosing and imaging MUC1-expressing
XX CC cancer cells and tissues, for purifying or isolating non-glycosylated,
XX CC underglycosylated or cancer-associated forms of MUC1 or MUC1 epitope-
XX CC containing molecules, and for therapeutically or prophylactically
XX CC treating cancer. The present sequence is human b1vph1-IL-2
XX CC immunocytokine protein. b1vph1 is mucin specific binding portion.
XX
XX SQ Sequence 381 AA;

Query Match 100.0%; Score 644; DB 22; Length 381;
Best Local Similarity 100.0%; Pred. No. 9.2e-50;
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QVOLVSGGLVQPGSLRLSCAASGFTFRSNAMGWVRQAPGKLEWVSGISGGSTYY 60
DB 1 qvqlvsgggglvpggslrlscaasgftfrsnamgwvrqapkglewsgisggstyy 60
QY 61 ADSVKGRFTISRDNKNTLYLQMSLRAEDTAVYYCAKHTGGGVWDPIDYWGQGLTVTVS 120
DB 61 adsvkgrftisrdsnknltlylqmslraedtavyycahkhtgggvwdpidywgqglvtvts 120
QY 121 s 121
DB 121 s 121

RESULT 3
AAE12715
ID AAE12715 standard; Protein: 451 AA.
XX
XX AC AAE12715;
XX
XX DT 04-JAN-2002 (first entry)
XX
XX DE Human recombinant immunoglobulin (Ig) heavy-chain region.
XX
XX KW Human; tumour-associated antigen mucin-1; MUC-1; adenocarcinoma;
XX KW heavy chain region; cancer; breast; ovary; lung; bladder;

```

KW cytotatic; therapy; immunoglobulin; Ig.  
 XX Homo sapiens.  
 XX WO200175110-A2.  
 XX 11-OCT-2001.  
 XX 30-MAR-2001; 2001WO-US0589.  
 XX 30-MAR-2000; 2000US-0538913.  
 XX (DYAX-) DYAX CORP.  
 XX Hoogenboom HRJM, Henderikx MPG;  
 XX WPI: 2001-626437/72.  
 XX N-PSDB; AAD20745.  
 XX Novel isolated tumor-associated antigen mucin-1-specific binding member  
 PT for diagnosing and treating cancer, comprises mucin-1 binding domain or  
 PT its portion for binding to an epitope of the protein core of mucin-1  
 XX  
 PS Claim 12: Page 106-108; 126pp; English.  
 XX The invention relates to an isolated tumour-associated antigen mucin-1  
 CC (MUC-1)-specific binding member comprising an antigen binding domain  
 CC region having an antibody variable light (VL) or heavy (VH) region,  
 CC or a complementarity determining region (CDR) of VL or VH. MUC1-specific  
 CC binding member is useful for diagnosing cancer, preferably adenocarcinoma  
 CC The binding of MUC1-specific binding member to MUC1 is detected by a  
 CC detection method selected from enzyme-linked immunosorbent assay,  
 CC magnetic resonance imaging, scintillation counting, and x-ray film.  
 CC MUC1-specific binding member is useful for treating cancer, preferably  
 CC adenocarcinoma, in an individual, where the cancer is present in tissue  
 CC of the breast, ovary, lung, or bladder of the individual. MUC1-specific  
 CC binding member is useful for diagnosing and imaging MUC1-expressing  
 CC cancer cells and tissues, for purifying or isolating non-glycosylated,  
 CC underglycosylated or cancer-associated forms of MUC1 or MUC1 epitope-  
 CC containing molecules, and for therapeutically or prophylactically  
 CC treating cancer. The present sequence is human recombinant immunoglobulin  
 CC (Ig) heavy chain region (variable VH and CH constant heavy chain).  
 XX  
 SQ Sequence 451 AA:  
 Query Match 100.0%; Score 644; DB 22; Length 451;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-49;  
 Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 QYQLVSGGLVPGGSLRLSCAASGTFTRSNANGWYRQAPGKLEWVSGISGGSTYY 60  
 Db 1 QYQLVSGGGLVPGGSLRLSCAASGTFTRSNANGWYRQAPGKLEWVSGISGGSTYY 60  
 QY 61 ADSVGRFTISRDNSKNTLYLQMNSLRAEDTAVYCAKHTGGGVWVPIDYWGQGLTLVTS 120  
 Db 61 adsvgrftisrdnsknltlylqmslraedtavyccakhtgggvwpdywgqgltlvts 120  
 QY 121 S 121  
 Db 121 s 121  
 RESULT 4  
 AAE07029  
 ID AAE07029 standard; Protein: 121 AA.  
 XX AAE07029;  
 XX 16-OCT-2001 (first entry)  
 XX Human heavy chain variable (VH) region, VH clone 41.  
 DE  
 XX

KW Human; humanised antibody; CC-chemokine receptor; CCR2; nephrotropic;  
 KW neuroprotective; immunosuppressive; human immunodeficiency virus;  
 KW HIV infection; cytostatic; vasotropic; leukocyte trafficking; allergy;  
 KW inflammatory disorder; autoimmune disorder; rheumatoid arthritis; shock;  
 KW multiple sclerosis; atherosclerosis; stenosis; restenosis; asthma;  
 KW anaphylaxis; malignancy; inflammation; stenosis; allograft rejection;  
 KW fibrotic disease; angioplasty; acquired immune deficiency syndrome;  
 KW AIDS; inflammatory glomerulopathy; vascular intervention; ID9 antibody;  
 KW neonatal hyperplasia; VH; heavy chain variable region.  
 XX Homo sapiens.  
 XX  
 OS  
 XX  
 Key Location/Qualifiers  
 FH Region 31..35  
 FT /label= CDR1  
 FT /note= "Complementarity determining region 1"  
 FT Region 50..66  
 FT /label= CDR2  
 FT /note= "Complementarity determining region 2"  
 FT Region 99..110  
 FT /label= CDR3  
 FT /note= "Complementarity determining region 3"  
 XX  
 XX WO200157226-A1.  
 XX  
 XX 09-AUG-2001.  
 XX  
 XX 02-FEB-2001; 2001WO-US03537.  
 XX  
 XX 03-FEB-2000; 2000US-0497625.  
 XX  
 XX (MILL-) MILLENNIUM PHARM INC.  
 XX  
 XX Larosa GJ, Horvath C, Newman W, Jones ST, O'Brien S, O'Keefe T;  
 XX WPI: 2001-488888/53.  
 XX  
 XX Humanized immunoglobulin for treating a CC-chemokine receptor  
 PT 2-mediated disorder in a patient, comprises a binding specificity for  
 PT CCR2, and a non-human antigen binding region and human immunoglobulin  
 PT  
 XX  
 PS Disclosure: Page 174-175; 183pp; English.  
 XX  
 CC The patent discloses a humanised antibody or its antigen-binding  
 CC fragment, having binding specificity for CC-chemokine receptor 2  
 CC (CCR2), comprising an antigen binding region of non-human origin  
 CC and at least a portion of an immunoglobulin of human origin. The  
 CC humanised antibodies are useful for inhibiting the interaction of  
 CC a cell expressing CCR2. They are useful for inhibiting or treating  
 CC HIV infection. The proteins of the invention are useful for inhibiting  
 CC leukocyte trafficking, for treating CCR2-mediated disorders such as  
 CC inflammatory disorder, autoimmune disorders such as rheumatoid  
 CC arthritis and multiple sclerosis, atherosclerosis and atherosclerosis,  
 CC and for inhibiting restenosis. They are useful in therapy or diagnosis,  
 CC and in the manufacture of a medicament for treating CCR2 mediated  
 CC disease. They are also useful for treating allergy, anaphylaxis,  
 CC malignancy, chronic and acute inflammation, histamine and IgE,  
 CC mediated allergic reaction, shock, stenosis, allograft rejection,  
 CC fibrotic disease, asthma, inflammatory glomerulopathies, acquired  
 CC immune deficiency syndrome (AIDS), restenosis associated with vascular  
 CC intervention, including angioplasty and/or stent placement in a mammal.  
 CC Humanised antibodies are also useful for inhibiting narrowing of the  
 CC lumen of a vessel in a mammal, and inhibiting neonatal hyperplasia of  
 CC a vessel in a mammal, preferably associated with vascular intervention.  
 CC The present sequence is human heavy chain variable (VH) region,  
 CC VH clone 41.  
 XX  
 SQ Sequence 121 AA:  
 Query Match 85.1%; Score 548; DB 22; Length 121;  
 Best Local Similarity 87.6%; Pred. No. 1e-41;

1

1771 S 1771

AAE07017



CC The present sequence is human heavy chain variable (VH) region, 034514.

XX

SQ Sequence 116 AA;

Query Match 84.1%; Score 541.5; DB 22; Length 116;  
Best Local Similarity 88.4%; Pred. No. 3.7e-41;  
Matches 107; Conservative 2; Mismatches 7; Indels 5; Gaps

QY 1 QVOLVSGGGLVPGGSRLRSLCAASGFTFRSNAMGVFQAAPCKLEWVSGISGGSTYY 60  
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
1 evqlvesgggvlvpggsrlrslcaasgtfssyamswwrqpkclewvsalsggsgstyy 60

QY 61 ADSVKGRFTISRDNKNTLYQMNSLRABDTAVYYCAKHGTGGVADPIDYWGQGILVTYS 120  
Db 61 adsvkgrftisrdnskntlyqmnsiraedtavyycakdkdgsq-v----ywggsgtlvtys 115

QY 121 S 121  
Db 116 s 116

RESULT 8  
AAWI3520  
ID AAWI3520 standard; protein; 131 AA.  
AC AAWI3520;  
XX  
DT 28-OCT-1997 (first entry)  
DE Anti-melanoma antibody heavy chain clone V13.  
XX Human; monoclonal antitumour antigen vaccine; peripheral blood lymphocyte;  
KW cancer; tumorigenesis; anticancer vaccine.  
KK Homo sapiens.  
OS  
FH Key Location/Qualifiers  
FT Region 31..35  
FT /label= CDR1  
FT Region 50..66  
FT /label= CDR2  
FT Region 99..105  
FT /label= CDR3  
PN WO9702479-A2.  
PD 23-JAN-1997.  
XX  
PF 28-JUN-1996; 96WO-IB01032.  
PR 30-JUN-1995; 95US-0497647.  
PA (UYFA ) UNIV YALE.  
XX Cai X, Garen A;  
XX WPI; 1997-109061/10.  
XX  
PT Prodn. of human monoclonal anti-tumour antibodies - by screening a  
PT fusion phase library produced using peripheral blood lymphocytes  
PT from a cancer patient  
XX  
PS Claim 19; Page 58; 82pp; English.  
XX

A process for isolating and synthesising human monoclonal anti-tumour antibodies has been produced. The process involves: (a) constructing at least one fusion phase library from the peripheral blood lymphocytes (PBLs) of a cancer patient; (b) screening for anti-tumour antibodies in the phase library in a binding assay with cultured tumour cells of the same type as the patient's tumour; (c) removing extraneous antibodies by absorption against normal human cells; (d) cloning the phase selected in step (b) and (c); (e) assaying the specificity of the cloned phage by

```
CC incubating the phage with at least two types of cultured normal cells;
CC and (f) further testing the specificity of cloned phage that do not bind
CC to either cell line of cultured normal cells in further binding assays;
CC to cultured tumour cells derived from more than one other tumour that is
CC not the patient's tumour. The present sequence represents a human heavy
CC chain antibody, from an scFv antibody fusion phage library, produced by
CC a method as described above. The antibodies produced can be used for
CC diagnostic and therapeutic applications and for isolating tumour
CC antigens for studying tumourigenesis or for use as anti-cancer vaccines.
CC The human antibodies have low immunogenicity in humans compared to
CC murine monoclonal antibodies (MABs). Since the antibodies are isolated
CC from fusion phage libraries, their affinity and specificity for a
CC tumour cell line can be improved by genetic manipulations.
XX
SQ Sequence 131 AA;

Query Match 83.5%; Score 537.5; DB 18; Length 131;
Best Local Similarity 89.3%; Pred. No. 9.5e-41;
Matches 108; Conservative 1; Mismatches 7; Indels 5; Gaps 2;

QY 1 QVQLVQSGGGLVQPGGSLRLSCAASGFTFRSNAMGWVRQAPGKGLEWVSGTSGGGSTYY 60
Db 1 evqlvsgggglvqpggslrlscaasgftfssyamswwrqapgkglewvsaisgsggstyy 60

QY 61 ADSVKGRFTISRDNKNTLYQMNSLRAEDTAVYYCAKHTGGGVWDPTDYNWGQGLTVTVS 120
Db 61 adsvkgrftisrdnskntlyqmnsiraedtavyycak----gv-afdywgggtptvts 115

QY 121 S 121
Db 116 s 116

RESULT 9
AAEL2061
ID AAEL2061 standard; Protein; 117 AA.
AC AAEL2061;
XX
XX 18-DEC-2001 (first entry)
DT
DE Human anti-tissue factor subgroup III variable heavy chain (VH) domain.
XX
XX Human; tissue factor; TF; thrombotic disorder; variable heavy chain;
XX VH; anticoagulant; thrombolytic; cytostatic; antibacterial; vasotropic;
XX immunosuppressive; cerebroprotective; vaccine; thrombosis; septicemia;
XX tumour; metastasis; arteriosclerosis; restenosis; angioplasty; stroke;
XX inflammation; adult respiratory distress syndrome; ARDS; hypotension;
XX disseminated intravascular coagulopathy; DIC.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
FH Region 26..35
FT /label= CDR_H1
FT Region 50..66
FT /label= CDR_H2
FT Region 99..106
FT /label= CDR_H3
XX
XX WO200170984-A2.
XX
XX 27-SEP-2001.
PD
XX 08-MAR-2001; 2001WO-US07501.
XX
XX 16-MAR-2000; 2000US-189775P.
XX
XX (GETH ) GENENTECH INC.
XX
XX Kirchofer DK, Lowe DG, Presta LG;
XX
XX
```

```
DR WPI; 2001-616406/71.
XX
XX Identifying anti-tissue factor antibodies for treating thrombotic
PT disorders, comprises epitope mapping of anti-TF antibodies and
PT selecting antibodies that bind to C-terminal macro-molecular
PT substrate-binding region of TF -
XX
XX Disclosure; Fig 8; 75pp; English.
PS
XX The invention relates to a method for identifying anti-tissue factor
XX (anti-TF) antibodies with enhanced anticoagulant potency. The method
CC comprises subjecting a number of anti-TF antibodies to epitope mapping,
CC and selecting the antibodies that bind to an epitope comprising at least
CC a part of the C-terminal macromolecular substrate-binding region of
CC tissue factor (TF). The method is useful for identifying anti-TF
CC antibodies which are useful for blocking a TF-FVIIa associated process
CC or event and for treating a TF-VIIa related disease or disorder
CC e.g. thrombotic or coagulopathic disorder including deep venous
CC thrombosis, arterial thrombosis, stroke, tumour, metastasis,
CC thrombolysis, arteriosclerosis and restenosis following angioplasty,
CC acute and chronic indications such as inflammation, septic shock,
CC septicemia, hypotension, adult respiratory distress syndrome (ARDS) and
CC disseminated intravascular coagulopathy (DIC). The present sequence is
CC human anti-tissue factor subgroup III variable heavy chain (VH) domain.
XX
XX Sequence 117 AA;

Query Match 83.4%; Score 537; DB 22; Length 117;
Best Local Similarity 86.8%; Pred. No. 9.4e-41;
Matches 105; Conservative 4; Mismatches 8; Indels 4; Gaps 1;

QY 1 QVQLVQSGGGLVQPGGSLRLSCAASGFTFRSNAMGWVRQAPGKGLEWVSGTSGGGSTYY 60
Db 1 evqlvsgggglvqpggslrlscaasgftfssyamswwrqapgkglewvsgdggstyy 60

QY 61 ADSVKGRFTISRDNKNTLYQMNSLRAEDTAVYYCAKHTGGGVWDPTDYNWGQGLTVTVS 120
Db 61 adsvkgrftisrdnskntlyqmnsiraedtavyycarggg-----sdywgggtlvtvs 116

QY 121 S 121
Db 117 s 117

RESULT 10
AAU14320
ID AAU14320 standard; Protein; 313 AA.
XX
XX AAU14320;
XX
XX 24-OCT-2001 (first entry)
DT
DE Human novel protein #191.
XX
XX Human; novel protein; Antianemic; osteopathic; anti-inflammatory;
XX immunomodulatory; cytostatic; neuroprotective; vitamin; nootropic;
XX anticonvulsant; antiarthritic; cerebroprotective; antifungal; antiviral;
XX antibacterial; anti-allergic; dermatological; haemostatic; antiasthmatic;
XX thrombolytic; immunogen; antibody; gene therapy; neurological disorder;
XX Parkinson's disease; inflammatory disorder; cancer; asthma; osteoporosis;
XX tissue regeneration; immune disorder.
XX
XX Homo sapiens.
XX
XX WO200155437-A2.
XX
XX 02-AUG-2001.
PD
XX 25-JAN-2001; 2001WO-US02623.
XX
XX 25-JAN-2000; 2000US-0491404.
XX
XX
```

PA (HYSE-) HYSEQ INC.  
 XX Tang YT, Liu C, Drmanac RT;  
 XX WPI: 2001-451939/48.  
 DR N-PSDB: AAS22625.  
 XX Isolated polypeptides useful for treating anti-inflammatory diseases,  
 PT nervous system disorders, and for regenerating bone and cartilage -  
 XX Example 4; Page 630-631; 894pp; English.  
 XX The invention relates to polynucleotides encoding novel human  
 CC proteins or their active domains. The polypeptides, polynucleotides and  
 CC antibodies raised against the polypeptides are used in a method of  
 CC treatment of a mammal and prevention of disorders caused by the aberrant  
 CC protein expression or activity. The polypeptides can be used as  
 CC molecular weight markers, food supplements, and in antibody production.  
 CC The polypeptides are used to identify compounds which bind to the  
 CC polypeptides. Polynucleotides of the invention are used as probes and  
 CC primers, for sequencing, for chromosome or gene mapping, in the  
 CC production of recombinant proteins, and in generating anti-sense DNA or  
 CC RNA and in gene therapy. Polypeptides of the invention can be used to  
 CC target drugs to a tumour, in assays to determine biological activity, to  
 CC raise antibodies/elicit an immune response, to determine quantitative  
 CC protein levels, as tissue markers, and to isolate receptors or ligands.  
 CC Polypeptides of the invention may also be useful in treating platelet  
 CC disorders, stem cell disorders, regenerating bone, cartilage, tendon,  
 CC ligament and/or nerve tissue, wound healing, treating burns, promoting  
 CC the proliferation, differentiation and survival of stem cells, as a  
 CC contraceptive, treating osteoporosis and osteoarthritis, anaemia,  
 CC Alzheimer's, Parkinson's and Huntington's diseases, amyotrophic lateral  
 CC sclerosis, stroke, immune deficiencies resulting from bacterial, viral or  
 CC fungal infection or from autoimmunity, cancer, allergy, asthma,  
 CC graft-versus-host disease, eczema, haemophilia, thrombosis,  
 CC anti-inflammatory diseases, nervous system disorders, and infection.  
 CC The present sequence represents a protein of the invention.  
 XX  
 XX Sequence 313 AA;

Query Match 83.4%; Score 537; DB 22; Length 313;  
 Best Local Similarity 83.5%; Pred. No. 2.7e-40;  
 Matches 106; Conservative 4; Mismatches 11; Indels 6; Gaps 1;  
 QY 1 QVQLVQSGGGLVQPQGGSLRLSCAASGFTFRSNAMGWVRQAPGKGLRWVSGISGSGSTYY 60  
 :||||:|||||  
 Db 61 evqllesggglvqpqgsirlscaasgftfssyamsvrqapqgkglewsaisgsggstyy 120  
 QY 61 ADSVKGRFTISRDNKNTLYLQMSLRRAEDTAVYYCAKHTGGGVWDP-----IDYWGQG 114  
 :||||:|||||  
 Db 121 adsvkgrftisrdnsknltlylqmslrdaedtavycakshpgyydydsggsyfydywggg 180  
 QY 115 TLVTYSS 121  
 :|||||  
 Db 181 tlvtvss 187

RESULT 11  
 AAY02472  
 ID AAY02472 standard; Protein; 240 AA.

XX AAY02472;  
 AC AAY02472;  
 XX 15-JUL-1999 (first entry)

XX A single chain antibody (ScFv).

XX Screening; functional polypeptide; ligand; non-functional;  
 KW enrichment; single chain antibody; ScFv.  
 XX Unidentified.

OS  
 PN  
 XX

PN WO9920749-A1.  
 XX 29-APR-1999.  
 XX 20-OCT-1998; 98WO-GB03135.  
 XX 21-NOV-1997; 97US-0066729.  
 PR 20-OCT-1997; 97GB-0022131.  
 PR 13-NOV-1997; 97US-0065428.  
 XX (MEDI-) MEDICAL RES COUNCIL.  
 XX Tomlinson I, Winter G;  
 PI WPI: 1999-288302/24.  
 DR N-PSDB: AAX36070.  
 XX Screening for functional polypeptides which bind a ligand  
 PT Disclosure; Fig 2; 67pp; English.  
 XX The specification describes a method for screening for functional  
 CC polypeptides which bind a ligand. The method comprises contacting a  
 CC repertoire of polypeptides with a generic ligand, and then screening  
 CC selected functional polypeptides with a target ligand. The method  
 CC permits the removal from a chosen repertoire of polypeptides those  
 CC which are non-functional, e.g. as a result of the introduction of  
 CC frame-shift mutations, stop codons, folding mutants or expression  
 CC mutants which would be or are incapable of binding to any target  
 CC ligand. The method also permits the enrichment of a chosen repertoire  
 CC of polypeptides for those polypeptides which are functional, well folded  
 CC and highly expressed. The polypeptides obtained can be used in  
 CC diagnostic, prophylactic and therapeutic procedures. The present  
 CC sequence represents the single chain antibody (ScFv) that forms the  
 CC basis of a library according to the invention.  
 XX  
 XX Sequence 240 AA;

Query Match 83.3%; Score 536.5; DB 20; Length 240;  
 Best Local Similarity 86.8%; Pred. No. 2.2e-40;  
 Matches 105; Conservative 3; Mismatches 8; Indels 5; Gaps 1;  
 QY 1 QVQLVQSGGGLVQPQGGSLRLSCAASGFTFRSNAMGWVRQAPGKGLRWVSGISGSGSTYY 60  
 :||||:|||||  
 Db 1 evqllesggglvqpqgsirlscaasgftfssyamsvrqapqgkglewsaisgsggstyy 60  
 QY 61 ADSVKGRFTISRDNKNTLYLQMSLRRAEDTAVYYCAKHTGGGVWDPIDYWGQGLTIVTS 120  
 :||||:|||||  
 Db 61 adsvkgrftisrdnsknltlylqmslrdaedtavycakshpgyydydsggsyfydywgggtitvts 115  
 QY 121 S 121  
 :|||||  
 Db 116 s 116

RESULT 12  
 AAW13532  
 ID AAW13532 standard; protein; 117 AA.

XX AAW13532;  
 AC AAW13532;  
 XX 28-OCT-1997 (first entry)

XX Anti-melanoma antibody C55 from VH antibody fusion phage library.

XX Human; monoclonal antitumour antibody; peripheral blood lymphocyte;  
 KW cancer; tumourigenesis; anticancer vaccine.  
 XX Homo sapiens.

OS  
 PN  
 XX

WO9702479-A2.



XX OS Synthetic.  
OS Homo sapiens.  
XX FH Location/Qualifiers  
FT Region 31..35 "Heavy chain variable region (VH) complementarity  
FT /note= "determining region 1 (CDRI)"  
FT Region 50..66  
FT /note= "VH-CDR2"  
FT Region 99..108  
FT /note= "VH-CDR3"  
FT Region 157..170  
FT /note= "Light chain variable region (VL) complementarity  
FT /note= "determining region 1 (CDRI)"  
FT Region 186..192  
FT /note= "VL-CDR2"  
FT Region 225..235  
FT /note= "VL-CDR3"  
XX WO9955367-A1.  
XX 04-NOV-1999.  
XX 23-APR-1999; 99WO-US07395.  
XX 24-APR-1998; 98US-0082953.  
XX 12-FEB-1999; 99US-0250056.  
XX (REGC ) UNIV CALIFORNIA.  
XX Marks JD, Poul MA;  
XX WPI: 2000-072168/06.  
XX N-PSDB; AA258235.  
XX Novel internalizing antibodies used to treat cancer cells -  
XX Claim 3; Page 81; 85pp; English.  
XX This sequence represents an internalising humanised antibody,  
XX scFv F5, which specifically binds to the extracellular domain  
XX of the c-erbB-2 receptor, the protein product of the HER/neu oncogene.  
XX The scFv F5 antibody binds to the epitope of the c-erbB-2 receptor that  
XX is bound by F5 antibodies. On binding the c-erbB-2 receptor, the  
XX antibody is transported into the cell. The c-erbB-2 receptor is a marker  
XX protein which is overexpressed by 30-50% of breast carcinomas and other  
XX adenocarcinomas, and thus provides a useful cell surface marker for  
XX specifically targeting tumour cells. The antibodies of the invention  
XX are used as tumour-targeting molecules for diagnosis and treatment. The  
XX antibodies can be attached to effector molecules. The effector molecules  
XX may include cytotoxins such as ricin, abrin or Pseudomonas exotoxin;  
XX radionuclides; ligands such as growth factors; therapeutic agents such  
XX as vinblastine, vindesine or melphalan; ribozymes; or antisense  
XX molecules. The antibodies may also be used for in vivo or in vitro  
XX detection and/or quantitation of the c-erbB-2 receptor and thus diagnosis  
XX and/or localisation of cancers characterised by expression of c-erbB-2.  
XX Although antibodies have previously been used to target tumour cells,  
XX their success has been limited. The utility of prior art antibodies has  
XX been hampered by the paucity of tumour specific antibodies, antibody  
XX immunogenicity, low binding affinity, and poor tumour penetration.  
XX Immunogenicity could be avoided and toxicity reduced if high affinity  
XX tumour specific human antibodies were available. However, the production  
XX of human monoclonal antibodies using conventional hybridoma technology  
XX has proven difficult. Also, most of the antibodies produced react with  
XX antigens that are also common to non-malignant cells, which makes them  
XX unsuitable for use as tumour-targeting molecules. The antibodies of the  
XX invention overcome these difficulties, as they are targeted to a  
XX tumour-specific antigen, and avoid the problem of immunogenicity as they  
XX are human in origin.  
XX Sequence 246 AA:  
SQ

Query Match 82.8%; Score 533; DB 21; Length 246;  
Best Local Similarity 86.8%; Pred. No. 4.7e-40;  
Matches 105; Conservative 2; Mismatches 12; Indels 2; Gaps 1;  
QY 1 QYOLVOSGGGLVOPGSLRLSCAASGFTFSNANGWVRQAPGK:LEWVSGLSGSGSTYY 60  
DB 1 qvqlvesggglvqpqgsrlscaasgftfsyamswrqapgk:kvysaisrgidntcy 60  
QY 61 ADSVKRGFTISRDNKNTLYLQMSLRAEDTAVYCAKHTGGGVWV:PIDYWGQGLTVTS 120  
DB 61 adsvkgrftisrdnsntlylqmslraedtaavycakmtsnat:afdywgqgltlvtvs 118  
QY 121 S 121  
DB 119 S 119  
RESULT 15  
AAB75051  
ID AAB75051 standard; Protein: 227 AA.  
XX AC AAB75051;  
XX 19-JUL-2001 (first entry)  
XX TRO005 HuMab Happa chain protein sequence 1CC6H.  
XX Human; antibody; immunoglobulin; interleukin 8; 103; immunogen;  
KW human antibody phage display library; immunisation; transgenic animal.  
XX Homo sapiens.  
XX Synthetic.  
XX WO200125492-A1.  
XX 12-APR-2001.  
XX 02-OCT-2000; 2000WO-US27237.  
XX 02-OCT-1999; 99US-0157415.  
XX 01-DEC-1999; 99US-0453234.  
XX (BIOS-) BIOSITE DIAGNOSTICS INC.  
XX (GENP-) GENPHARM INT SUBSIDIARY OF MEDAREX INC.  
XX Buechler J, Valkirs G, Gray J, Lonberg N;  
XX WPI: 2001-335567/35.  
XX Producing a human antibody phage display library comprises providing a  
XX transgenic animal whose genome comprises human immunoglobulin genes and  
XX isolating nucleic acids encoding antibody chains from lymphatic cells -  
XX Example 37; Page 122-123; 161pp; English.  
XX The present invention describes a method (M1) for producing a human  
XX antibody phage display library (L1), comprising:  
XX (1) providing a nonhuman  
XX transgenic animal (II) whose genome comprises human immunoglobulin genes;  
XX (2) isolating nucleic acids encoding human antibody chains (III) from  
XX lymphatic cells; and (3) forming a library of display packages whose  
XX members comprise a nucleic acid encoding (III) which is displayed from  
XX the package. The method is used for producing a human antibody display  
XX library, e.g., a Fab phage display library. The display method may be  
XX used to screen nucleic acids encoding antibody chains obtained from  
XX immunised nonhuman transgenic animals, and from this a population of  
XX antibodies may be prepared. Production of a human monoclonal antibodies  
XX display library using this method means there is no need to immunise  
XX humans with antigens, and the difficulties faced with immortalising B  
XX cells are avoided. AAB75051 to AAB75056 and AAB75054 to AAB75056  
XX represent sequences used in the exemplification of the present invention.  
XX Sequence 227 AA:  
SQ

Query Match 82.6%; Score 532; DB 22; Length 227;  
Best Local Similarity 86.8%; Pred. No. 5.3e-40;  
Matches 105; Conservative 3; Mismatches 11; Indels 2; Gaps 1;

Qy 1 QVQLVQSGGGLVOPGGSLRLSCAASGFTPRSNAMGWVROAPGKGLEWVSGISGGSTYY 60  
Db :|||||  
Qy 1 evqlvqsggglvqpggsrlriscasgftfsnyamswvrqpgkglewsainygggstyy 60  
Db :|||||

Qy 61 ADSVKGRFTTISRDNKNTLYLQMNLSRAEDTAVYYCAKHTGGGVWDPIDYWGQGLTVTVS 120  
Db :|||||

Qy 61 adsvkgrftlisdnsktlylqmnsdraedtavyycakhmrvgl--fdywggtlvtvs 118  
Db :|||||

Qy 121 S 121  
Db 119 s 119

Search completed: October 4, 2002, 10:21:47  
Job time: 265 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 4, 2002, 10:27:38 ; Search time 24.71 Seconds  
(without alignments)  
470.531 Million cell updates/sec

Title: US-09-822-698A-3

Perfect score: 121

Sequence: 1 QVOLVSGGGLVQPGSLRL.....GGVWDPIYWGQGLVTVSS 121

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 283138 seqs, 96089334 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : PIR\_71.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	63	52.1	120	2 S36278	Ig heavy chain V r
2	51	42.1	90	2 S24248	Ig heavy chain V r
3	48	39.7	98	2 S26889	Ig heavy chain V r
4	48	39.7	99	2 S24259	Ig heavy chain V r
5	48	39.7	100	2 S24258	Ig heavy chain V r
6	48	39.7	101	2 S24257	Ig heavy chain V r
7	48	39.7	102	2 S24260	Ig heavy chain V r
8	48	39.7	104	2 S24255	Ig heavy chain V r
9	48	39.7	106	2 S24256	Ig heavy chain V r
10	48	39.7	108	2 S24257	Ig heavy chain V r
11	48	39.7	109	2 PH1648	Ig heavy chain V r
12	48	39.7	112	2 PH1647	Ig heavy chain V r
13	48	39.7	117	2 A45953	Ig heavy chain pre
14	48	39.7	118	2 S31121	Ig heavy chain - h
15	48	39.7	119	2 C36005	Ig heavy chain V r
16	48	39.7	119	2 D36005	Ig heavy chain V r
17	48	39.7	119	2 S31108	Ig heavy chain - h
18	48	39.7	119	2 S31108	Ig heavy chain - h
19	48	39.7	120	2 S48798	Ig heavy chain V r
20	48	39.7	123	2 S31114	Ig heavy chain - h
21	48	39.7	127	2 S38489	Ig heavy chain - h
22	48	39.7	138	2 S31666	Ig heavy chain V r
23	48	39.7	160	2 S05271	Ig heavy chain pre
24	47	38.8	105	2 S24249	Ig heavy chain V r
25	47	38.8	121	2 S31113	Ig heavy chain - h
26	47	38.8	121	2 S56673	Ig heavy chain - h
27	44	36.4	97	2 S26935	Ig heavy chain V r
28	44	36.4	97	2 S46462	Ig heavy chain V r
29	44	36.4	113	2 S24247	Ig heavy chain V r

30 44 36.4 116 2 B28966 Ig heavy chain pre  
31 42 34.7 117 2 A34964 Ig heavy chain pre  
32 42 34.7 124 2 S20782 Ig heavy chain V r  
33 40 33.1 98 2 PL0116 Ig heavy chain V-I  
34 40 33.1 98 2 S29543 Ig heavy chain V r  
35 40 33.1 111 2 PH1643 Ig heavy chain V r  
36 40 33.1 121 2 S19666 Ig heavy chain V r  
37 40 33.1 122 2 S31119 Ig heavy chain - h  
38 40 33.1 130 2 PL0098 Ig heavy chain pre  
39 40 33.1 134 2 S31688 Ig heavy chain V r  
40 40 33.1 135 2 S31598 Ig heavy chain V r  
41 40 33.1 139 2 S31674 Ig heavy chain V r  
42 40 33.1 147 2 I37780 Ig variable region  
43 39 32.2 98 2 S29546 Ig heavy chain V r  
44 39 32.2 108 2 PH1642 Ig heavy chain V r  
45 39 32.2 109 2 PH1644 Ig heavy chain V r

#### ALIGNMENTS

RESULT 1 50-66

S36278  
Ig heavy chain V region (clone alpha-THY-23) - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 03-Feb-1994 #sequence\_revision 03-Feb-1994 #text\_change 23-Jul-1999  
C:Accession: S36278  
R:Griffiths, A.D.; Malmqvist, M.; Marks, J.D.; Bye, J.M.; Embleton, M.J.; McCafferty, E.M.O. J. 12, 725-734, 1993  
A:Title: Human anti-self antibodies with high specificity from phage display library  
A:Reference number: S36256; MUID:93178448  
A:Accession: S36278  
A:Status: preliminary; nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-120 <R1>  
A:Cross-references: EMBL:X18830; NID:g33114; PIDN:CAA79282.1; PID:g939894  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotetramer; immunoglobulin  
F:15-98/Domain: immunoglobulin homology <IMM>

Query Match 52.1%; Score 63; DB 2; Length 120;  
Best Local Similarity 100.0%; Pred. NO. 1.9e-55;  
Matches 63; Conservative 0; Mismatches 0; Gaps 0;

QY 36 WVRQAPGKLEWVSGISGSGSTYYADSVKGRFTISRDNKNTLYIQMNSLRADTAVYY 95  
|||||

Db 36 WVRQAPGKLEWVSGISGSGSTYYADSVKGRFTISRDNKNTLYIQMNSLRADTAVYY 95  
|||||

QY 96 CAK 98  
|||

Db 96 CAK 98

RESULT 2

S24248  
Ig heavy chain V region (VH26) - human

C:Species: Homo sapiens (man)

C:Date: 19-Feb-1994 #sequence\_revision 10-Nov-1995 #text\_change 23-Jul-1999

C:Accession: S24248

R:Stewart, A.K.; Huang, C.; Stollar, B.D.; Schwartz, R.S.

A:Description: A single VH gene predominates in the rearranged and expressed human B

A:Reference number: S24247

A:Accession: S24248

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-90 <STE>

A:Cross-references: EMBL:X67069; NID:g38395; PIDN:CAA47454.1; PID:g38396

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

Query Match 42.1%; Score 51; DB 2; Length 90;  
Best Local Similarity 100.0%; Pred. No. 1.4e-43;  
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 41 PGKLEWVSGISGGSTYYADSVKGRFTISRDNKNTLYLQMSLRAEDT 91  
|||||  
Db 40 PGKLEWVSGISGGSTYYADSVKGRFTISRDNKNTLYLQMSLRAEDT 90

## RESULT 3

S26889  
Ig heavy chain V region (DP-47) - human  
C:Species: Homo sapiens (man)  
C:Date: 22-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 23-Jul-1999  
C:Accession: S26889  
R:Tomlinson, I.M.; Walter, G.; Marks, J.D.; Llewelyn, M.B.; Winter, G.  
J. Mol. Biol. 227, 776-798, 1992  
A:Title: The repertoire of human germline V(H) sequences reveals about fifty groups of V  
A:Reference number: S26889; MUID:9302117  
A:Accession: S26889  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-98 <TOM>  
A:Cross-references: EMBL:Z12347; NID:g32914; PIDN:CAA78217.1; PID:g32915  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotetramer; immunoglobulin  
F:15-98/Domain: immunoglobulin homology <IMM>

Query Match 39.7%; Score 48; DB 2; Length 98;  
Best Local Similarity 100.0%; Pred. No. 1.5e-40;  
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 ISGSGSTYYADSVKGRFTISRDNKNTLYLQMSLRAEDTAVYYCAK 98  
|||||  
Db 51 ISGSGSTYYADSVKGRFTISRDNKNTLYLQMSLRAEDTAVYYCAK 98

## RESULT 4

S24259  
Ig heavy chain V region (VH26-DK1-JH4) - human  
C:Species: Homo sapiens (man)  
C:Date: 19-Feb-1994 #sequence\_revision 10-Nov-1995 #text\_change 23-Jul-1999  
C:Accession: S24259  
R:Stewart, A.K.; Huang, C.; Stollar, B.D.; Schwartz, R.S.  
submitted to the EMBL Data Library, June 1992  
A:Description: A single VH gene predominates in the rearranged and expressed human B cell  
A:Reference number: S24247  
A:Accession: S24259  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-99 <STE>  
A:Cross-references: EMBL:X67067; NID:g38391; PIDN:CAA47452.1; PID:g38392  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotetramer; immunoglobulin  
F:1-83/Domain: immunoglobulin homology <IMM>

Query Match 39.7%; Score 48; DB 2; Length 99;  
Best Local Similarity 100.0%; Pred. No. 1.5e-40;  
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 ISGSGSTYYADSVKGRFTISRDNKNTLYLQMSLRAEDTAVYYCAK 98  
|||||  
Db 36 ISGSGSTYYADSVKGRFTISRDNKNTLYLQMSLRAEDTAVYYCAK 83

## RESULT 5

S24258  
Ig heavy chain V region (VH26-DXP1-JH4) - human  
C:Species: Homo sapiens (man)  
C:Date: 19-Feb-1994 #sequence\_revision 10-Nov-1995 #text\_change 23-Jul-1999  
C:Accession: S24258

R:Stewart, A.K.; Huang, C.; Stollar, B.D.; Schwartz, R.S.  
submitted to the EMBL Data Library, June 1992  
A:Description: A single VH gene predominates in the rearranged and expressed human B  
A:Reference number: S24247  
A:Accession: S24258  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-100 <STE>  
A:Cross-references: EMBL:X67066; NID:g38389; PIDN:CAA47451.1; PID:g38390  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotetramer; immunoglobulin  
F:4-86/Domain: immunoglobulin homology <IMM>

Query Match 39.7%; Score 48; DB 2; Length 100;  
Best Local Similarity 100.0%; Pred. No. 1.5e-40;  
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 ISGSGSTYYADSVKGRFTISRDNKNTLYLQMSLRAEDTAVYYCAK 98  
|||||  
Db 39 ISGSGSTYYADSVKGRFTISRDNKNTLYLQMSLRAEDTAVYYCAK 86

## RESULT 6

S24257  
Ig heavy chain V region (VH26-DXP1-JH4) - human  
C:Species: Homo sapiens (man)  
C:Date: 19-Feb-1994 #sequence\_revision 10-Nov-1995 #text\_change 21-Jan-2000  
C:Accession: S24257  
R:Stewart, A.K.; Huang, C.; Stollar, B.D.; Schwartz, R.S.  
submitted to the EMBL Data Library, June 1992  
A:Description: A single VH gene predominates in the rearranged and expressed human B  
A:Reference number: S24247  
A:Accession: S24257  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-101 <STE>  
A:Cross-references: EMBL:X67065; NID:g38387; PIDN:CAA47450.1; PID:g38388  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotetramer; immunoglobulin  
F:11-93/Domain: immunoglobulin homology <IMM>

Query Match 39.7%; Score 48; DB 2; Length 101;  
Best Local Similarity 100.0%; Pred. No. 1.5e-40;  
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 ISGSGSTYYADSVKGRFTISRDNKNTLYLQMSLRAEDTAVYYCAK 98  
|||||  
Db 46 ISGSGSTYYADSVKGRFTISRDNKNTLYLQMSLRAEDTAVYYCAK 93

## RESULT 7

S24260  
Ig heavy chain V region (VH26-DAL1-JH4) - human  
C:Species: Homo sapiens (man)  
C:Date: 19-Feb-1994 #sequence\_revision 10-Nov-1995 #text\_change 06-Jun-1997  
C:Accession: S24260  
R:Stewart, A.K.; Huang, C.; Stollar, B.D.; Schwartz, R.S.  
submitted to the EMBL Data Library, June 1992  
A:Description: A single VH gene predominates in the rearranged and expressed human B  
A:Reference number: S24247  
A:Accession: S24260  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-102 <STE>  
A:Cross-references: EMBL:X67068  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotetramer; immunoglobulin  
F:5-87/Domain: immunoglobulin homology <IMM>

Query Match 39.7%; Score 48; DB 2; Length 102;



Best Local Similarity 100.0%; Pred. No. 1.6e-40;  
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 ISGSGSTYYADSVKGRFTISRDNKNTLYLQMSLRAEDTAVYYCAK 98  
|||||  
Db 40 ISGSGSTYYADSVKGRFTISRDNKNTLYLQMSLRAEDTAVYYCAK 87  
|||||

## RESULT 8

Ig heavy chain V region (VH26-DLR5-JH4) - human  
C:Species: Homo sapiens (man)  
C:Date: 19-Feb-1994 #sequence\_revision 10-Nov-1995 #text\_change 06-Jun-1997  
C:Accession: S24255  
R:Stewart, A.K.; Huang, C.; Stollar, B.D.; Schwartz, R.S.  
submitted to the EMBL Data Library, June 1992  
A:Description: A single VH gene predominates in the rearranged and expressed human B cell  
A:Reference number: S24247  
A:Accession: S24255  
A:Status: Preliminary  
A:Molecule type: DNA  
A:Residues: 1-104 <STE>  
A:Cross-references: EMBL:X67063  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotetramer; immunoglobulin  
F:2-84/Domain: immunoglobulin homology <IMM>

Query Match 39.7%; Score 48; DB 2; Length 104;  
Best Local Similarity 100.0%; Pred. No. 1.6e-40;  
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 ISGSGSTYYADSVKGRFTISRDNKNTLYLQMSLRAEDTAVYYCAK 98  
|||||  
Db 37 ISGSGSTYYADSVKGRFTISRDNKNTLYLQMSLRAEDTAVYYCAK 84  
|||||

## RESULT 9

Ig heavy chain V region (VH26-DXP4-JH6) - human  
C:Species: Homo sapiens (man)  
C:Date: 19-Feb-1994 #sequence\_revision 10-Nov-1995 #text\_change 23-Jul-1999  
C:Accession: S24256  
R:Stewart, A.K.; Huang, C.; Stollar, B.D.; Schwartz, R.S.  
submitted to the EMBL Data Library, June 1992  
A:Description: A single VH gene predominates in the rearranged and expressed human B cell  
A:Reference number: S24247  
A:Accession: S24256  
A:Status: Preliminary  
A:Molecule type: DNA  
A:Residues: 1-106 <STE>  
A:Cross-references: EMBL:X67064; NID:G38385; PID:CAA7449.1; PID:G38386  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotetramer; immunoglobulin  
F:4-86/Domain: immunoglobulin homology <IMM>

Query Match 39.7%; Score 48; DB 2; Length 106;  
Best Local Similarity 100.0%; Pred. No. 1.6e-40;  
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 ISGSGSTYYADSVKGRFTISRDNKNTLYLQMSLRAEDTAVYYCAK 98  
|||||  
Db 39 ISGSGSTYYADSVKGRFTISRDNKNTLYLQMSLRAEDTAVYYCAK 86  
|||||

## RESULT 10

Ig heavy chain V region (clone 2B8) - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 24-Feb-1994 #sequence\_revision 24-Feb-1994 #text\_change 16-Aug-1996  
C:Accession: PH1648  
R: Hillson, J.L.; Karr, N.S.; Opplinger, I.R.; Mannik, M.; Sasso, E.H.

J. Exp. Med. 178, 331-336, 1993  
A:Title: The structural basis of germline-encoded VH3 immunoglobulin binding to staph  
A:Reference number: PH1642; MUID:93301610  
A:Accession: PH1648  
A:Molecule type: mRNA  
A:Residues: 1-108 <HIL>  
A:Experimental source: B cell  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotetramer; immunoglobulin  
F:7-90/Domain: immunoglobulin homology <IMM>

Query Match 39.7%; Score 48; DB 2; Length 108;  
Best Local Similarity 100.0%; Pred. No. 1.6e-40;  
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 ISGSGSTYYADSVKGRFTISRDNKNTLYLQMSLRAEDTAVYYCAK 98  
|||||  
Db 43 ISGSGSTYYADSVKGRFTISRDNKNTLYLQMSLRAEDTAVYYCAK 90  
|||||

## RESULT 11

PH1649  
Ig heavy chain V region (clone 2E7) - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 24-Feb-1994 #sequence\_revision 24-Feb-1994 #text\_change 16-Aug-1996  
C:Accession: PH1649  
R: Hillson, J.L.; Karr, N.S.; Opplinger, I.R.; Mannik, M.; Sasso, E.H.  
J. Exp. Med. 178, 331-336, 1993  
A:Title: The structural basis of germline-encoded VH3 immunoglobulin binding to staph  
A:Reference number: PH1642; MUID:93301610  
A:Accession: PH1649  
A:Molecule type: mRNA  
A:Residues: 1-109 <HIL>  
A:Experimental source: B cell  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotetramer; immunoglobulin  
F:7-90/Domain: immunoglobulin homology <IMM>

Query Match 39.7%; Score 48; DB 2; Length 109;  
Best Local Similarity 100.0%; Pred. No. 1.7e-40;  
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 ISGSGSTYYADSVKGRFTISRDNKNTLYLQMSLRAEDTAVYYCAK 98  
|||||  
Db 43 ISGSGSTYYADSVKGRFTISRDNKNTLYLQMSLRAEDTAVYYCAK 90  
|||||

## RESULT 12

PH1647  
Ig heavy chain V region (clone 2D10) - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 24-Feb-1994 #sequence\_revision 24-Feb-1994 #text\_change 16-Aug-1996  
C:Accession: PH1647  
R: Hillson, J.L.; Karr, N.S.; Opplinger, I.R.; Mannik, M.; Sasso, E.H.  
J. Exp. Med. 178, 331-336, 1993  
A:Title: The structural basis of germline-encoded VH3 immunoglobulin binding to staph  
A:Reference number: PH1642; MUID:93301610  
A:Accession: PH1647  
A:Molecule type: mRNA  
A:Residues: 1-112 <HIL>  
A:Experimental source: B cell  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotetramer; immunoglobulin  
F:7-90/Domain: immunoglobulin homology <IMM>

Query Match 39.7%; Score 48; DB 2; Length 112;  
Best Local Similarity 100.0%; Pred. No. 1.7e-40;  
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 ISGSGSTYYADSVKGRFTISRDNKNTLYLQMSLRAEDTAVYYCAK 98  
|||||  
Db 43 ISGSGSTYYADSVKGRFTISRDNKNTLYLQMSLRAEDTAVYYCAK 90  
|||||

|||||  
Db 43 ISGSGSTYYADSVKGRFTISRDNKNTLYLQMNSLRAEDTAVYYCAK 90  
RESULT 13  
A45953  
Ig heavy chain precursor V-III region (VH26) - human  
C:Species: Homo sapiens (man)  
C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 07-May-1999  
C:Accession: A45953  
R:Chen, P.P.; Liu, M.F.; Sinha, S.; Carson, D.A.  
Arthritis Rheum. 31, 1429-1431, 1988  
A:Title: A 16/6 idiotype-positive anti-DNA antibody is encoded by a conserved V-H gene  
C:Reference number: A45953; MUID:89050363  
A:Accession: A45953  
A:Status: preliminary; not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 1-117 <CHE>  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotetramer; immunoglobulin  
F:15-98/Domain: immunoglobulin homology <IMM>

Query Match 39.7%; Score 48; DB 2; Length 117;  
Best Local Similarity 100.0%; Pred. No. 1.8e-40;  
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 ISGSGSTYYADSVKGRFTISRDNKNTLYLQMNSLRAEDTAVYYCAK 98  
|||||  
Db 70 ISGSGSTYYADSVKGRFTISRDNKNTLYLQMNSLRAEDTAVYYCAK 117

## RESULT 14

S31121  
Ig heavy chain - human  
C:Species: Homo sapiens (man)  
C:Date: 02-Dec-1993 #sequence\_revision 26-May-1995 #text\_change 17-Mar-1999  
C:Accession: S31121  
R:Raaphorst, F.M.; Timmers, E.; Kenter, M.J.H.; van Tol, M.J.D.; Vossen, J.M.; Schuurman  
Eur. J. Immunol. 22, 247-251, 1992  
A:Title: Restricted utilization of germ-line V(H)3 genes and short diverse third complement  
A:Reference number: S31104; MUID:92111633  
A:Accession: S31121  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: mRNA  
A:Residues: 1-118 <RAA>  
A:Cross-references: EMBL:X62971  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1991  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotetramer; immunoglobulin  
F:15-98/Domain: immunoglobulin homology <IMM>

Query Match 39.7%; Score 48; DB 2; Length 118;  
Best Local Similarity 100.0%; Pred. No. 1.8e-40;  
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 ISGSGSTYYADSVKGRFTISRDNKNTLYLQMNSLRAEDTAVYYCAK 98  
|||||  
Db 51 ISGSGSTYYADSVKGRFTISRDNKNTLYLQMNSLRAEDTAVYYCAK 98

## RESULT 15

C36005  
Ig heavy chain V region (30pl) - human  
C:Species: Homo sapiens (man)  
C:Date: 21-Dec-1990 #sequence\_revision 21-Dec-1990 #text\_change 16-Aug-1996  
C:Accession: C36005  
R:Schroeder Jr., H.W.; Wang, J.Y.  
Proc. Natl. Acad. Sci. U.S.A. 87, 6146-6150, 1990  
A:Title: Preferential utilization of conserved immunoglobulin heavy chain variable gene  
A:Reference number: A36005; MUID:90349571  
A:Accession: C36005

Query Match 39.7%; Score 48; DB 2; Length 119;  
Best Local Similarity 100.0%; Pred. No. 1.8e-40;  
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 ISGSGSTYYADSVKGRFTISRDNKNTLYLQMNSLRAEDTAVYYCAK 98  
|||||  
Db 51 ISGSGSTYYADSVKGRFTISRDNKNTLYLQMNSLRAEDTAVYYCAK 98

Search completed: October 4, 2002, 10:30:40  
Job time: 182 sec



100

101

GenCore version 4.5  
 Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 4, 2002, 10:29:48 ; Search time 43.48 Seconds  
 (without alignments)  
 481.426 Million cell updates/sec

Title: US-09-822-698A-3  
 Perfect score: 121  
 Sequence: 1 QVQLVSGGSLVQPGGSLRL.....GGVWDPIYWGQGLIVTVSS 121

Scoring table: OLIGO  
 Gapop 60.0 , Gapext 60.0

Searched: 562222 seqs, 172994929 residues

Word size : 0

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SPTREMBL\_19:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phase:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_virus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a  
 score greater than or equal to the score of the result being printed,  
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	33.1	113	4 Q9UL90	Q9ul90 homo sapien
2	38	31.4	112	4 Q9UGP3	Q9ugp3 homo sapien
3	34	28.1	116	4 Q9UL93	Q9ul93 homo sapien
4	34	28.1	118	4 Q9UL72	Q9ul72 homo sapien
5	25	20.7	597	4 Q96BB9	Q96bb9 homo sapien
6	23	19.0	95	4 Q9ULB6	Q9ulb6 homo sapien
7	23	19.0	118	4 Q9UL91	Q9ul91 homo sapien
8	23	19.0	121	4 Q9UL71	Q9ul71 homo sapien
9	21	17.4	147	4 Q9I509	Q9y509 homo sapien
10	19	15.7	112	4 Q9HCC1	Q9hcc1 homo sapien
11	18	14.9	122	4 Q9UL84	Q9ul84 homo sapien
12	17	14.0	104	4 Q9UL87	Q9ul87 homo sapien
13	16	13.2	131	4 Q9UL88	Q9ul88 homo sapien
14	14	11.6	480	11 Q91XE1	Q91xel mus musculu
15	14	11.6	494	4 Q96K68	Q96k68 homo sapien
16	14	11.6	496	4 Q96KX8	Q96kx8 homo sapien

17	13	10.7	119	4 Q9UL73	Q9ul73 homo sapien
18	13	10.7	119	11 Q920E7	Q920e7 mus musculu
19	13	10.7	121	4 Q9UL96	Q9ul96 homo sapien
20	13	10.7	122	4 Q9UL75	Q9ul75 homo sapien
21	13	10.7	124	4 Q9UL92	Q9ul92 homo sapien
22	13	10.7	384	4 Q9UF60	Q9uf60 homo sapien
23	12	9.9	124	6 Q9N0W6	Q9n0w4 oryctolagus
24	12	9.9	124	6 Q9N0W4	Q9n0w4 oryctolagus
25	12	9.9	298	11 Q9QYF0	Q9qyf0 mus musculu
26	12	9.9	488	11 Q91WR1	Q91wr1 mus musculu
27	12	9.9	613	4 Q96EV0	Q96ev0 homo sapien
28	11	9.1	64	11 Q61750	Q61750 mus musculu
29	11	9.1	102	11 Q9UL79	Q9ul79 mus musculu
30	11	9.1	110	11 Q9UL83	Q9ul83 mus musculu
31	11	9.1	117	11 Q9ZIC6	Q9zic6 mus musculu
32	11	9.1	119	4 Q9UL94	Q9ul94 homo sapien
33	11	9.1	120	11 Q920E8	Q920e8 mus musculu
34	11	9.1	134	6 P83049	P83049 sus scrofa
35	11	9.1	139	11 Q924R5	Q924r5 mus musculu
36	11	9.1	463	11 Q9ULC4	Q9ulc4 mus musculu
37	11	9.1	473	11 Q9DBL4	Q9db14 mus musculu
38	11	9.1	473	11 Q91Z05	Q91z05 mus musculu
39	11	9.1	487	11 Q99KA4	Q99ka4 mus musculu
40	11	9.1	496	4 Q96DK0	Q96dk0 homo sapien
41	10	8.3	150	4 Q9Y298	Q9y298 homo sapien
42	10	8.3	416	4 Q9NPP6	Q9npp6 homo sapien
43	10	8.3	437	11 Q9RIA4	Q9ria4 mus musculu
44	10	8.3	479	11 Q99M22	Q99m22 mus musculu
45	10	8.3	479	11 Q91WP5	Q91wp5 mus musculu

ALIGNMENTS

RESULT	1
Q9UL90	PRELIMINARY; PRT; 113 AA.
ID	Q9UL90;
AC	Q9UL90;
DT	01-MAY-2000 (TREMBlrel. 13, Created)
DT	01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT	01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE	MYOSIN-REACTIVE IMMUNOGLOBULIN HEAVY CHAIN VARIANTS REGION
DE	(FRAGMENT).
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Euthera; Primates; Catarrhini; Homnida; Homo.
OX	NCBI_TaxID=9606;
RN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=98277139; PubMed=9614934;
RA	Wu X., Liu B., Van der Merwe P.L., Kallis N.N., Kennedy S.M.,
RA	Young D.C.;
RT	"Myosin-reactive autoantibodies in rheumatic carditis and normal fetus.";
RL	Clin. Immunol. Immunopathol. 87:184-192(1998).
DR	EMBL: AF035024; AAD56260.1;
DR	HSSP: P01772; 2PBA.
DR	InterPro: IPR003006; IG_MHC.
DR	InterPro: IPR003596; IG_V.
DR	Pfam: PF00047; IG; 1.
DR	SMART: SM00406; IGV; 1.
FT	NON_TER 1
FT	NON_TER 113
SQ	SEQUENCE 113 AA; 12437 MW; ED57FDD19086D07F CRC64;

Query Match 33.1%; Score 40; DB 4; Length 113;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-33;

Matches 40; Conservative 0; Mismatches 0; Indels 0; Caps 0;

QY 59 YYADSVKGFETSRNSKNTLYLQMSLRADTAVYYCAK

DB 59 YYADSVKGFETSRNSKNTLYLQMSLRADTAVYYCAK 98

```

RESULT 2
Q9UGP3 ID Q9UGP3 PRELIMINARY; PRT; 112 AA.
AC Q9UGP3;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE MYOSIN-REACTIVE IMMUNOGLOBULIN HEAVY CHAIN (FRAGMENT).
GN IGH.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Zafiroopoulos A., Kandilogianaki M., Dahlenborg C., Borrebaeck C.A.K.,
RA Krambovitis E.;
RT "Induction of somatic mutations in human B cells by in vitro
RT immunization.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ132560; CAB5078.1; -.
DR HSSP; P01772; 2FB4.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; Igv; 1.
DR NON_TER 1
FT NON_TER 112
SQ SEQUENCE 112 AA; 12236 MW; 0A1F17082C7A8CE3 CRC64;

Query Match 31.4%; Score 38; DB 4; Length 112;
Best Local Similarity 100.0%; Pred. No. 2.1e-31;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 YADSVKGRFTISRDNSKNTLYQMNSLRADTA 97
Db 43 YADSVKGRFTISRDNSKNTLYQMNSLRADTA 80

RESULT 3
Q9UL93 ID Q9UL93 PRELIMINARY; PRT; 116 AA.
AC Q9UL93;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE MYOSIN-REACTIVE IMMUNOGLOBULIN HEAVY CHAIN VARIABLE REGION
DE (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Wu X., Liu B., Van der Merwe P.L., Kallis N.N., Berney S.M.,
RA Young D.C.;
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
RT fetus.";
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL; AF035021; AAD56257.1; -.
DR HSSP; P01772; 2FB4.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; Igv; 1.
DR NON_TER 1
FT NON_TER 116
SQ SEQUENCE 116 AA; 12434 MW; 0DA0348154DD6061 CRC64;

Query Match 31.4%; Score 38; DB 4; Length 112;
Best Local Similarity 100.0%; Pred. No. 2.1e-31;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 YADSVKGRFTISRDNSKNTLYQMNSLRADTA 97
Db 43 YADSVKGRFTISRDNSKNTLYQMNSLRADTA 80

RESULT 4
Q9UL72 ID Q9UL72 PRELIMINARY; PRT; 118 AA.
AC Q9UL72;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE MYOSIN-REACTIVE IMMUNOGLOBULIN HEAVY CHAIN VARIABLE REGION
DE (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Wu X., Liu B., Van der Merwe P.L., Kallis N.N., Berney S.M.,
RA Young D.C.;
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
RT fetus.";
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL; AF035042; AAD56278.1; -.
DR HSSP; P01772; 2FB4.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; Igv; 1.
DR NON_TER 1
FT NON_TER 118
SQ SEQUENCE 118 AA; 12872 MW; B4D1A5944B2D5CCA CRC64;

Query Match 28.1%; Score 34; DB 4; Length 118;
Best Local Similarity 100.0%; Pred. No. 2.9e-27;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 59 YADSVKGRFTISRDNSKNTLYQMNSLRADTA 92
Db 58 YADSVKGRFTISRDNSKNTLYQMNSLRADTA 91

RESULT 5
Q96BB9 ID Q96BB9 PRELIMINARY; PRT; 597 AA.
AC Q96BB9;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE HYPOTHETICAL 65.0 KDA PROTEIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE=PRIMARY B-CELLS FROM TONSILS;
RA Strausberg R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC015760; AAHL5760.1; -.
DR Hypothetical protein.
KW Hypothetical protein.
SQ SEQUENCE 597 AA; 65039 MW; 4FCA3AD8ECE263D9 CRC64;

```

```

Query Match 28.1%; Score 34; DB 4; Length 116;
Best Local Similarity 100.0%; Pred. No. 2.9e-27;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 59 YADSVKGRFTISRDNSKNTLYQMNSLRADTA 92
Db 58 YADSVKGRFTISRDNSKNTLYQMNSLRADTA 91

RESULT 4
Q9UL72 ID Q9UL72 PRELIMINARY; PRT; 118 AA.
AC Q9UL72;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE MYOSIN-REACTIVE IMMUNOGLOBULIN HEAVY CHAIN VARIABLE REGION
DE (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Wu X., Liu B., Van der Merwe P.L., Kallis N.N., Berney S.M.,
RA Young D.C.;
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
RT fetus.";
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL; AF035042; AAD56278.1; -.
DR HSSP; P01772; 2FB4.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; Igv; 1.
DR NON_TER 1
FT NON_TER 118
SQ SEQUENCE 118 AA; 12872 MW; B4D1A5944B2D5CCA CRC64;

Query Match 28.1%; Score 34; DB 4; Length 118;
Best Local Similarity 100.0%; Pred. No. 2.9e-27;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 59 YADSVKGRFTISRDNSKNTLYQMNSLRADTA 92
Db 58 YADSVKGRFTISRDNSKNTLYQMNSLRADTA 91

RESULT 5
Q96BB9 ID Q96BB9 PRELIMINARY; PRT; 597 AA.
AC Q96BB9;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE HYPOTHETICAL 65.0 KDA PROTEIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE=PRIMARY B-CELLS FROM TONSILS;
RA Strausberg R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC015760; AAHL5760.1; -.
DR Hypothetical protein.
KW Hypothetical protein.
SQ SEQUENCE 597 AA; 65039 MW; 4FCA3AD8ECE263D9 CRC64;

```

```
Query Match      20.7%; Score 25; DB 4; Length 597;
Best Local Similarity 100.0%; Pred. No. 2.3e-17;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 ISGGSTYYADSVKGRFTISRDN 75
      |||||
Db 70 ISGGSTYYADSVKGRFTISRDN 94
      |||||

RESULT 6
Q9ULB6 PRELIMINARY; PRT; 95 AA.
AC Q9ULB6;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE IMUNOGLOBULIN HEAVY CHAIN (FRAGMENT).
GN VH.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Tange Y., Kavano H.;
RT "Human VH gene sequence.";
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB035268; BAA87067.1; -.
DR HSSP; P01772; 2FB4.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; Igv; 1.
DR NON_TER 1
FT NON_TER 95
SQ SEQUENCE 95 AA; 10527 MW; 90A8C6D16D22574A CRC64;

Query Match      19.0%; Score 23; DB 4; Length 95;
Best Local Similarity 100.0%; Pred. No. 5.8e-16;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 SGGGLVQPGSLRLSCAASGTF 29
      |||||
Db 6 SGGGLVQPGSLRLSCAASGTF 28
      |||||

RESULT 7
Q9UL91 PRELIMINARY; PRT; 118 AA.
AC Q9UL91;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE MYOSIN-REACTIVE IMMUNOGLOBULIN HEAVY CHAIN VARIABLE REGION
  (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Wu X., Liu B., Van der Merwe P.L., Kallis N.N., Berney S.M.,
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
  fetus.";
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL; AF035023; AAD56259.1; -.
DR HSSP; P01772; 2FB4.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; Igv; 1.
DR NON_TER 1
FT NON_TER 95
SQ SEQUENCE 95 AA; 10527 MW; 90A8C6D16D22574A CRC64;

Query Match      19.0%; Score 23; DB 4; Length 121;
Best Local Similarity 100.0%; Pred. No. 7.1e-16;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 55 GGSTYYADSVKGRFTISRDN 77
      |||||
Db 55 GGSTYYADSVKGRFTISRDN 77
      |||||

RESULT 9
Q9Y509 PRELIMINARY; PRT; 147 AA.
ID Q9Y509;
AC Q9Y509;
DT 01-NOV-1999 (TReMBLrel. 12, Created)
DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE VH3 PROTEIN (FRAGMENT).
GN VH3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Cao J., Vescio R.A., Rettig M.B., Hong C.H., Kim J.C.,
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
  fetus.";
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL; AF035023; AAD56259.1; -.
DR HSSP; P01772; 2FB4.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
SQ SEQUENCE 121 AA; 13154 MW; 2F045CCFASD50736 CRC64;
```

RA Lichtenstein A.K., Berenson J.R.;  
 RT "A CD10-positive subset of malignant cells is identified in multiple  
 RL myeloma using PCR with patient-specific immunoglobulin gene primers.";  
 RL Leukemia 9:1948-1953(1995).  
 DR EMBL: S80860; AAD14339.1; -.  
 DR HSSP: P01772; 2FB4.  
 DR InterPro: IPR003006; Ig\_MHC.  
 DR InterPro: IPR003596; Ig\_V.  
 DR Pfam: PF00047; Ig; 1.  
 DR SMART: SM00406; IGV; 1.  
 FT NON\_TER 1  
 SQ SEQUENCE 147 AA; 15768 MW; 8489FCAA7BC925C CRC64;

Query Match 17.4%; Score 21; DB 4; Length 147;  
 Best Local Similarity 100.0%; Pred. No. 9.9e-14;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 63 SVKGFETISRDNSKNTLYLQM 83  
 |||||  
 DB 63 SVKGFETISRDNSKNTLYLQM 83

RESULT 10  
 Q9HCC1 PRELIMINARY; PRT; 112 AA.  
 AC Q9HCC1  
 DT 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE SINGLE CHAIN FV (FRAGMENT).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Kikuchi M., Takeda C., Tsujimoto Y., Asada S., Nagata K.;  
 RT "An antibody fragment2A3 specific for native lysozyme: isolation from a  
 RL human synthetic phage display library and characterization.";  
 RL submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AB049915; BAB16829.1; -.  
 DR HSSP: P01772; 2FB4.  
 DR InterPro: IPR003599; Ig.  
 DR InterPro: IPR003600; Ig\_Like.  
 DR InterPro: IPR003006; Ig\_MHC.  
 DR InterPro: IPR003596; Ig\_V.  
 DR Pfam: PF00047; Ig; 1.  
 DR SMART: SM00409; IGV; 1.  
 DR SMART: SM00410; IGV; 1.  
 DR SMART: SM00410; IGV; 1.  
 FT NON\_TER 1  
 FT NON\_TER 112  
 SQ SEQUENCE 112 AA; 12243 MW; 24F1A45EC3B84788 CRC64;

Query Match 15.7%; Score 19; DB 4; Length 112;  
 Best Local Similarity 100.0%; Pred. No. 9.1e-12;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 79 LYQMSLRADTAVTYCA 97  
 |||||  
 DB 79 LYQMSLRADTAVTYCA 97

RESULT 11  
 Q9UL84 PRELIMINARY; PRT; 122 AA.  
 AC Q9UL84  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE MYOSIN-REACTIVE IMMUNOGLOBULIN HEAVY CHAIN VARIABLE REGION

DE (FRAGMENT).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE:98277139; PubMed=9614934;  
 RA Wu X., Liu B., Van der Merwe P.L., Kallis N.N., Berney S.M.,  
 RA Young D.C.;  
 RT "Myosin-reactive autoantibodies in rheumatic carditis and normal  
 RT fetus.";  
 RL Clin. Immunol. Immunopathol. 87:184-192(1998).  
 DR EMBL: AF035027; AAD56263.1; -.  
 DR HSSP: P01772; 2FB4.  
 DR InterPro: IPR003006; Ig\_MHC.  
 DR InterPro: IPR003596; Ig\_V.  
 DR Pfam: PF00047; Ig; 1.  
 DR SMART: SM00406; IGV; 1.  
 DR SMART: SM00406; IGV; 1.  
 FT NON\_TER 1  
 FT NON\_TER 122  
 SQ SEQUENCE 122 AA; 13579 MW; 36054D41366545B8 CRC64;

Query Match 14.9%; Score 18; DB 4; Length 122;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-10;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 LQMSLRADTAVTYCAK 98  
 |||||  
 DB 81 LQMSLRADTAVTYCAK 98

RESULT 12  
 Q9UL87 PRELIMINARY; PRT; 104 AA.  
 ID Q9UL87  
 AC Q9UL87  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE MYOSIN-REACTIVE IMMUNOGLOBULIN HEAVY CHAIN VARIABLE REGION  
 DE (FRAGMENT).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE:98277139; PubMed=9614934;  
 RA Wu X., Liu B., Van der Merwe P.L., Kallis N.N., Berney S.M.,  
 RA Young D.C.;  
 RT "Myosin-reactive autoantibodies in rheumatic carditis and normal  
 RT fetus.";  
 RL Clin. Immunol. Immunopathol. 87:184-192(1998).  
 DR EMBL: AF035027; AAD56263.1; -.  
 DR HSSP: P01810; 2FBJ.  
 DR InterPro: IPR003006; Ig\_MHC.  
 DR InterPro: IPR003596; Ig\_V.  
 DR Pfam: PF00047; Ig; 1.  
 DR SMART: SM00406; IGV; 1.  
 DR SMART: SM00406; IGV; 1.  
 FT NON\_TER 1  
 FT NON\_TER 104  
 SQ SEQUENCE 104 AA; 11598 MW; 611D1A3F40E96E7B CRC64;

Query Match 14.0%; Score 17; DB 4; Length 104;  
 Best Local Similarity 100.0%; Pred. No. 9.9e-10;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 WVRQAPGKGLWVSGIS 52  
 |||||  
 DB 15 WVRQAPGKGLWVSGIS 31



```

RESULT 15
Q96K68
ID Q96K68 PRELIMINARY; PRT; 494 AA.
AC Q96K68;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
CDNA FJ1714473 FTS, CLONE MAMMA1001080, HIGHLY SIMILAR TO HOMO
SAPIENS SNC73 PROTEIN (SNC73) MRNA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homini; Homo.
NCBI_TaxID=9606;
[1]
RN SEQUENCE FROM N.A.
RC TISSUE=MAMMARY GLAND;
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuuchi T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
RA Wagatsuma M., Hosoi T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ochi Y., Takiguchi S.,
RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
RA Yanamoto J., Wakamatsu A., Nakamura Y., Nagahara Y., Masuho Y.,
RA Ninomiya K., Iwayanagi T.;
RA "NEDO human cDNA sequencing project.";
RT Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AK027379; BAB55072.1;
SQ SEQUENCE 494 AA; 53088 MW; 9A1D7AEB5AE4C0E C1564;

Query Match 11.6%; Score 14; DB 4; Length 494;
Best Local Similarity 100.0%; Pred. No. 4.7e-06;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 WVRQAPGKGLEWVS 49
DB 55 WVRQAPGKGLEWVS 68
|||||
|||||

Search completed: October 4, 2002, 10:36:14
Job time: 386 sec

```

```

RESULT 13
Q9UL88 PRELIMINARY; PRT; 131 AA.
Q9UL88;
AC Q9UL88;
01-MAY-2000 (TREMBLrel. 13, Created)
01-MAY-2000 (TREMBLrel. 13, Last sequence update)
01-DEC-2001 (TREMBLrel. 19, Last annotation update)
MYOSIN-REACTIVE IMMUNOGLOBULIN HEAVY CHAIN VARIABLE REGION
(FRAGMENT)).
DE DE
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_Taxid=9606;
[1]
SEQUENCE FROM N.A.
MEDLINE=98277139; PubMed=9614934;
Wu X., Liu B., Van der Merwe P.L., Kalis N.N., Berney S.M.,
Young D.C.;
"MYOSIN-reactive autoantibodies in rheumatic carditis and normal
fetus";
Clin. Immunol. Immunopathol. 87:184-192(1998).
EMBL; AF035026; AAD56262.1; -.
HSSP; P01810; 2FBJ.
InterPro; IPR003006; Ig_MHC.
InterPro; IPR003596; Ig_v.
Pfam; PF000047; Ig_1.
DR SMART; SMC0406; IGV; 1.
DR SMART; 1
NON_TER 1
FT NON_TER 131
SEQUENCE 131 AA; 14142 MW; 96E7D668E375DEA0 CRC64;

Query Match 13.2%; Score 16; DB 4; Length 131;
Best Local Similarity 100.0%; Pred. No. 1.3e-08;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 PGGSRLRSLCAASGFTF 29
|||||
DB 14 PGGSRLRSLCAASGFTF 29
|||||

```

```

RESULT 14
Q91XEL PRELIMINARY; PRT; 480 AA.
ID Q91XEL
AC Q91XEL;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE UNKNOWN (PROTEIN FOR IMAGE:4224494) (FRAGMENT).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=COLON;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DDbj databases.
DR EMBL; BC010798; AAH10798.1; -.
FT NON_TER 1
SQ SEQUENCE 480 AA; 51936 MW; 20B9234EEF2B41ED CRC64;

Query Match 11.6%; Score 14; DB 11; Length 480;
Best Local Similarity 100.0%; Pred. No. 4.6e-06;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 PGSLRLSCAASGF 27
Db 32 PGSLRLSCAASGF 45

```



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 4, 2002, 10:25:48 ; Search time 50.76 seconds  
(without alignments)  
264.774 Million cell updates/sec

Title: US-09-822-698a-3

Perfect score: 121  
Sequence: 1 QVQLVSGGGLVPGGSLRL.....GCWDPIDYWGQGLTVTVSS 121

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 747574 seqs, 111073796 residues

Word size : 0

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : A\_Geneseq\_032802.\*

1:	/SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.*
2:	/SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.*
3:	/SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.*
4:	/SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.*
5:	/SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.*
6:	/SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.*
7:	/SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.*
8:	/SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.*
9:	/SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.*
10:	/SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.*
11:	/SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.*
12:	/SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.*
13:	/SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.*
14:	/SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.*
15:	/SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.*
16:	/SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.*
17:	/SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.*
18:	/SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.*
19:	/SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.*
20:	/SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.*
21:	/SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.*
22:	/SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	121	100.0	121	AAE12706	Human PH1 Fab anti
2	121	100.0	381	AAE12707	Human bivPH1-IL-2
3	121	100.0	451	AAE12715	Human recombinant
4	49	40.5	122	AAW78432	Antibody heavy chain
5	48	39.7	98	AAW40093	Anti-hiL12 antibody
6	48	39.7	98	AAV50972	Human FvIII antibody
7	48	39.7	98	AAE12710	Antibody variable
8	48	39.7	116	AAE07017	Human heavy chain
9	48	39.7	117	AAE66315	Human immunoglobulin
10	48	39.7	120	AAE07022	Human heavy chain
11	48	39.7	121	AAE07029	Human heavy chain

12	48	39.7	122	22	AAU02589	Anti-adipocyte mon
13	48	39.7	125	22	AAE07013	Human heavy chain
14	48	39.7	125	22	AAE07021	Human heavy chain
15	48	39.7	128	18	AAW06242	Heavy chain variab
16	48	39.7	128	22	AAE07014	Human heavy chain
17	48	39.7	131	18	AAW13520	Anti-melanoma anti
18	48	39.7	240	20	AAU02472	A single chain ant
19	48	39.7	313	22	AAU14320	Human novel protei
20	47	38.8	108	22	AAG55318	Anti-IL-18 antibod
21	47	38.8	112	21	AAV95184	Anti-platelet glyc
22	47	38.8	112	21	AAV95184	Anti-platelet glyc
23	47	38.8	113	19	AAW47179	Variable heavy (VH
24	47	38.8	115	21	AAV95210	Anti-platelet glyc
25	47	38.8	115	21	AAV95211	Anti-platelet glyc
26	47	38.8	118	18	AAW19878	CEA-specific antib
27	47	38.8	118	22	AAG65354	Anti-IL-18 antibod
28	47	38.8	120	18	AAW27553	Human Ab heavy cha
29	47	38.8	121	19	AAW47180	Variable heavy (VH
30	47	38.8	122	22	AAU02621	Anti-adipocyte mon
31	47	38.8	122	18	AAW13537	Anti-melanoma anti
32	47	38.8	124	22	AAE07026	Human heavy chain
33	47	38.8	130	19	AAW71048	scFv-B11 antibody
34	47	38.8	140	18	AAW13524	Anti-melanoma anti
35	47	38.8	235	22	AAG65320	Anti-IL-18 antibod
36	47	38.8	237	21	AAV95219	Anti-platelet glyc
37	47	38.8	240	21	AAV15124	Anti-human CTLA-4
38	47	38.8	240	21	AAV15125	Anti-murine CTLA-4
39	47	38.8	240	22	AAW46052	Human TF anti-idio
40	47	38.8	245	22	AAW67620	Human leukocyte an
41	47	38.8	281	18	AAW27560	Consensus single c
42	47	38.8	293	22	AAG65715	Amino acid sequenc
43	45	37.2	227	22	AAW75051	TRO005 HuMab Happa
44	44	36.4	95	18	AAW35098	Product of target
45	44	36.4	97	21	AAW40090	Anti-hiL12 antibody

#### ALIGNMENTS

RESULT 1  
AAE12706  
ID AAE12706 standard; Protein; 121 AA.

XX AAE12706;

DT 04-JAN-2002 (first entry)

XX Human: PH1 Fab antibody variable heavy chain region (VH).

XX Human; tumour-associated antigen mucin-1; MUC-1; adenocarcinoma; VH;  
variable heavy chain region; cancer; breast; ovary; lung; bladder;  
cytostatic; therapy; PH1 antibody.

XX Homo sapiens.

XX Location/Qualifiers

EH Key

FT Region

FT /label= CDR1

FT /note= "Complementarity determining region 1"

FT Region

FT /label= CDR2

FT /note= "Complementarity determining region 2"

FT Region

FT /label= FR3

FT /note= "Framework region 3"

FT Region

FT /note= "VH domain"

FT Region

FT /label= CDR3

FT /note= "Complementarity determining region 3"

FT Region

FT /label= FR4

FT /note= "Framework region 4"

XX WO200175110-A2.  
 XX 11-OCT-2001.  
 XX 30-MAR-2001; 2001WO-US10589.  
 XX 30-MAR-2000; 2000US-0538913.  
 XX (DYAX-) DYAX CORP.  
 XX Hoogenboom HRJM, Henderikx MPG;  
 XX WPI; 2001-626437/72.  
 XX N-PSDB; AAD20731.  
 XX Novel isolated tumor-associated antigen mucin-1-specific binding member  
 PT for diagnosing and treating cancer, comprises mucin-1 binding domain or  
 PT its portion for binding to an epitope of the protein core of mucin-1 -  
 XX  
 PS Claim 2; Page 94-95; 126pp; English.  
 XX The invention relates to an isolated tumor-associated antigen mucin-1  
 CC (MUC-1)-specific binding member comprising an antigen binding domain  
 CC region having an antibody variable light (VL) or heavy (VH) region,  
 CC or a complementarity determining region (CDR) of VL or VH. MUC1-specific  
 CC binding member is useful for diagnosing cancer, preferably adenocarcinoma  
 CC The binding of MUC1-specific binding member to MUC1 is detected by a  
 CC detection method selected from enzyme-linked immunosorbent assay,  
 CC magnetic resonance imaging, scintillation counting, and X-ray film.  
 CC MUC1-specific binding member is useful for treating cancer, preferably  
 CC adenocarcinoma, in an individual, where the cancer is present in tissue  
 CC of the breast, ovary, lung, or bladder of the individual. MUC1-specific  
 CC binding member is useful for diagnosing and imaging MUC1-expressing  
 CC cancer cells and tissues, for purifying or isolating non-glycosylated,  
 CC underglycosylated or cancer-associated forms of MUC1 or MUC1 epitope-  
 CC containing molecules, and for therapeutically or prophylactically  
 CC treating cancer. The present sequence is anti-MUC1 human Ph1 Fab  
 CC antibody VH region.  
 XX Sequence 121 AA;  
 SQ

Query Match 100.0%; Score 121; DB 22; Length 121;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-109;  
 Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 QVQLVSGGGLVOPGGSLRLSCAASGFTFRSNAMGWVRQAPGKLEWVSGISGSGSTYY 60  
 DB 1 qvqlvsgggllvqpgslrlscaasgftfrsnamgwrqapgkglewvsgisgsgstyy 60  
 QY 61 ADSVKGRFTISRDNKNTLYLQMSLRAEDTAVYYCAKHTGGVMDPIDYWGQGLTLTVTS 120  
 DB 61 adsvkgrftisrdnsknltlylqmslraedtavyycahkhtggvgwvdpdywgqgltlvts 120  
 QY 121 S 121  
 DB 121 s 121  
 RESULT 2  
 ID AAE12707 standard; Protein; 381 AA.  
 XX AAE12707;  
 XX 04-JAN-2002 (first entry)  
 XX Human b1vPH1-IL-2 immunocytokine protein.  
 XX Human; tumor-associated antigen mucin-1; MUC-1; adenocarcinoma; cancer;  
 KW breast; ovary; lung; bladder; cytostatic; therapy; immunocytokine.  
 XX

OS Homo sapiens.  
 XX WO200175110-A2.  
 XX 11-OCT-2001.  
 XX 30-MAR-2001; 2001WO-US10589.  
 XX 30-MAR-2000; 2000US-0538913.  
 XX (DYAX-) DYAX CORP.  
 XX Hoogenboom HRJM, Henderikx MPG;  
 XX WPI; 2001-626437/72.  
 XX N-PSDB; AAD20732.  
 XX Novel isolated tumor-associated antigen mucin-1-specific binding member  
 PT for diagnosing and treating cancer, comprises mucin-1 binding domain or  
 PT its portion for binding to an epitope of the protein core of mucin-1 -  
 XX  
 PS Claim 9; Page 95-97; 126pp; English.  
 XX The invention relates to an isolated tumor-associated antigen mucin-1  
 CC (MUC-1)-specific binding member comprising an antigen binding domain  
 CC region having an antibody variable light (VL) or heavy (VH) region,  
 CC or a complementarity determining region (CDR) of VL or VH. MUC1-specific  
 CC binding member is useful for diagnosing cancer, preferably adenocarcinoma  
 CC The binding of MUC1-specific binding member to MUC1 is detected by a  
 CC detection method selected from enzyme-linked immunosorbent assay,  
 CC magnetic resonance imaging, scintillation counting, and X-ray film.  
 CC MUC1-specific binding member is useful for treating cancer, preferably  
 CC adenocarcinoma, in an individual, where the cancer is present in tissue  
 CC of the breast, ovary, lung, or bladder of the individual. MUC1-specific  
 CC binding member is useful for diagnosing and imaging MUC1-expressing  
 CC cancer cells and tissues, for purifying or isolating non-glycosylated,  
 CC underglycosylated or cancer-associated forms of MUC1 or MUC1 epitope-  
 CC containing molecules, and for therapeutically or prophylactically  
 CC treating cancer. The present sequence is human b1vPH1-IL-2  
 CC immunocytokine protein. b1vPH1 is mucin specific binding portion.  
 XX Sequence 381 AA;  
 SQ

Query Match 100.0%; Score 121; DB 22; Length 381;  
 Best Local Similarity 100.0%; Pred. No. 3.2e-109;  
 Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 QVQLVSGGGLVOPGGSLRLSCAASGFTFRSNAMGWVRQAPGKLEWVSGISGSGSTYY 60  
 DB 1 qvqlvsgggllvqpgslrlscaasgftfrsnamgwrqapgkglewvsgisgsgstyy 60  
 QY 61 ADSVKGRFTISRDNKNTLYLQMSLRAEDTAVYYCAKHTGGVMDPIDYWGQGLTLTVTS 120  
 DB 61 adsvkgrftisrdnsknltlylqmslraedtavyycahkhtggvgwvdpdywgqgltlvts 120  
 QY 121 S 121  
 DB 121 s 121  
 RESULT 3  
 ID AAE12715 standard; Protein; 451 AA.  
 XX AAE12715;  
 XX 04-JAN-2002 (first entry)  
 XX Human recombinant immunoglobulin (Ig) heavy chain region.  
 XX Human; tumor-associated antigen mucin-1; MUC-1; adenocarcinoma;  
 KW heavy chain region; cancer; breast; ovary; lung; bladder;  
 XX

KW cytostatic; therapy; immunoglobulin; Ig.  
 XX Homo sapiens.  
 OS

PW W0200175110-A2.

PN 11-OCT-2001.

PP 30-MAR-2001; 2001WO-US10589.

PR 30-MAR-2000; 2000US-0538913.

PA (DYAX-) DYAX CORP.

XX Hoogenboom HRJM, Henderikx MPG;

PI WPI; 2001-626437/72.

XX N-PSDB; AAD20745.

DR Novel isolated tumor-associated antigen mucin-1-specific binding member  
 for diagnosing and treating cancer, comprises mucin-1 binding domain or  
 its portion for binding to an epitope of the protein core of mucin-1

PS Claim 12; Page 106-108; 126pp; English.

XX The invention relates to an isolated tumour-associated antigen mucin-1  
 CC (MUC-1)-specific binding member comprising an antigen binding domain  
 CC region having an antibody variable light (VL) or heavy (VH) region,  
 CC or a complementarity determining region (CDR) of VL or VH. MUC1-specific  
 CC binding member is useful for diagnosing cancer, preferably adenocarcinoma  
 CC The binding of MUC1-specific binding member to MUC1 is detected by a  
 CC detection method selected from enzyme-linked immunosorbent assay,  
 CC magnetic resonance imaging, scintillation counting, and X-ray film.  
 CC MUC1-specific binding member is useful for treating cancer, preferably  
 CC adenocarcinoma, in an individual, where the cancer is present in tissue  
 CC of the breast, ovary, lung, or bladder of the individual. MUC1-specific  
 CC binding member is useful for diagnosing and imaging MUC1-expressing  
 CC cancer cells and tissues, for purifying or isolating non-glycosylated,  
 CC underglycosylated or cancer-associated forms of MUC1 or MUC1 epitope-  
 CC containing molecules, and for therapeutically or prophylactically  
 CC treating cancer. The present sequence is human recombinant immunoglobulin  
 CC (Ig) heavy chain region (variable VH and CH constant heavy chain).

XX Sequence 451 AA;

Query Match 100.0%; Score 121; DB 22; Length 451;  
 Best Local Similarity 100.0%; Pred. No. 3.7e-109;  
 Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QVQLVQSGGGLVQPGGSLRLSCAASGFTFSRNAMGVROAPGKGLWVSGISGSGSTYY 60

Db 1 qvqlvqsgggivqpggslrlscaasgftfrsnamgvrwvrapgkglwvsgisgsgstyy 60

QY 61 ADSVKGRFTISRDNKNTLYLQMSLRADETAVYYCAKHTGGGVMDPIDYWGQGLPLTVS 120

Db 61 adsvkgrftisrdnskntlylqmslradetavyycahtgggvmdpidywgqglpltv 120

QY 121 S 121

Db 121 s 121

RESULT 4

AAW78432

ID AAW78432 standard; Protein; 122 AA.

XX AAW78432;

AC 11-MAY-1999 (first entry)

DT Antibody heavy chain targeted to HER3 clone 18.

DE

XX

KW Variant; antibody; heavy chain; light chain; immunoadhesin; immunoassay;  
 XX diagnosis; cancer; primer; PCR; amplification; diagnostic.  
 OS Synthetic.

PN W09850431-A2.

PP 12-NOV-1998.

PR 30-APR-1998; 98WO-US08762.

PR 24-JUN-1997; 97US-0050661.

PR 02-MAY-1997; 97US-0850058.

XX (GETH ) GENENTECH INC.

XX Arathoon R, Carter PJ, Merchant AM, Presta IG;

XX WPI; 1999-070091/06.

XX Selective preparation of multispecific antibodies with  
 PT heteromultimeric heavy chain and common light chain components,  
 PT useful for, e.g. in vivo diagnosis of cancer

PS Example 4; Fig 5; 69pp; English.

XX This sequence represents the heavy chain variable region for an antibody  
 CC that binds to the HER3 clone 18 protein. The sequence encoding the chain  
 CC is generated by a new method for preparing a multispecific Ab comprising  
 CC a first polypeptide (PP) and at least 1 extra PP, where: (i) the first PP  
 CC comprises a multimerisation domain (MD) forming an interface positioned  
 CC to interact with an interface of a MD of the extra PP; and (ii) the first  
 CC and extra PPs each have a binding domain, which comprises a heavy chain  
 CC and a light chain, where the variable light chains of the first and extra  
 CC PPs comprise a common sequence. The method comprises: (a) culturing a  
 CC host cell comprising nucleic acid encoding the first PP and extra PP, and  
 CC the variable light chain, such that the nucleic acid is expressed; and  
 CC (b) recovering the multispecific Ab from the culture. The method prepares  
 CC heteromultimeric PPs, such as bispecific Abs, bispecific immunoconjugates  
 CC and Ab-immunoconjugate chimeras. The method allows for the enhanced  
 CC formation of the desired heteromultimer relative to the undesired  
 CC heteromultimers and homomultimers. The Abs can be used in immunoassays  
 CC and for the in vitro or in vivo diagnosis of various diseases, such as  
 CC cancer.

XX Sequence 122 AA;

Query Match 40.5%; Score 49; DB 20; Length 122;  
 Best Local Similarity 100.0%; Pred. No. 8.4e-40;  
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 WVRQAPGKGLWVSGISGSGSTYYADSVKGRFTISRDNKNTLYLQMSLRAD 84

Db 36 wvrqapgkglwvsgisgsgstyyadvskgrftisrdnskntlylqmslr 84

RESULT 5

AAW40093

ID AAW40093 standard; Protein; 98 AA.

XX AAW40093;

XX 05-FEB-2001 (first entry)

DT Anti-hIL12 antibody H chain V region amino acid sequence SEQ ID 619.

DE Human; neutralising antibody; interleukin-12; IL-12; antiinflammatory;  
 XX complementarity determining region; CDR; antigenic; antiarthritic;  
 KW antisclerotic; neuroprotective; antiproliferative; antiasthmatic; cardiant;  
 KW antiparasitic; antibacterial; immunosuppressive; Crohn's disease;  
 KW multiple sclerosis; rheumatoid arthritis.

XX

```

OS Homo sapiens.
XX WO200056772-A1.
XX PD 28-SEP-2000.
XX PF 24-MAR-2000; 2000WO-US07946.
XX PR 25-MAR-1999; 99US-0126603.
XX PA (BADI) BASF AG.
XX PA (GENY) GENETICS INST INC.
XX PI Salfeld JG, Roguska M, Paskind M, Banerjee S, Tracey DE, White M;
XX PI Kaymakalan Z, Labkovsky B, Sakorafas P, Friedrich S, Myles A;
XX PI Veldman GM, Venturini A, Warne NW, Widom A, Elvin JG, Duncan AR;
XX PI Derbyshire EJ, Carmen S, Smith S, Holtet TL, Du Fou SL;
XX DR WPI; 2000-638250/61.
XX
XX New human antibody specific for human interleukin-12 (IL-12) used to
XX treat disorders characterized by aberrant IL-12 expression e.g. Crohn's
XX disease and multiple sclerosis -
XX
XX Claim 75: Page 121; 377pp; English.
XX
XX This invention relates to a new human antibody specific for human
XX interleukin-12 (IL-12). The invention also includes antigen binding
XX portions that bind to IL-12. Sequences AAB39485-B39516 represent human
XX anti-IL-12 antibody heavy and light chain complementarity determining
XX region (CDR) amino acid sequences, and also includes variable region
XX amino acid sequences. Other variable region amino acid sequences are
XX given in AAB39517-B39560 and AAB40068-B40149. Sequences AAB39561-B39771
XX represent anti-IL-12 CDR3 related amino acid sequences, AAB39772-B40063
XX represent other CDR sequences. Light chain CDR3 consensus sequences are
XX given in AAB40064-B40087. Primers used in the identification and
XX construction of the antibodies of the invention are given in
XX AAC61082-C61071. The antibody of the invention is a neutralizing
XX antibody and has antineoplastic; antiarthritic; antisclerotic;
XX antiinflammatory; neuroprotective; antiposrotatic; antiasthmatic;
XX cardiant; antiparasitic; antibacterial and immunosuppressive activity.
XX The antibodies or antigen-binding fragments are useful in the treatment
XX of disorders associated with detrimental release of human IL-12,
XX especially Crohn's disease, multiple sclerosis and rheumatoid arthritis.
XX They can also be used in the manufacture of a pharmaceutical composition
XX to treat human IL-12 disorders.
XX
XX Sequence 98 AA;

Query Match 39.7%; Score 48; DB 21; Length 98;
Best Local Similarity 100.0%; Pred. No. 6.5e-39;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 ISGSGGTYADSVKGRFTISRDNKNTLYLQMNSLRAEDTAVYYCAK 98
|||||
Db 51 isgsggtyyadvskgrftisrdnsknltlylqmnsltraedtavyycak 98
|||||

RESULT 6
AY50972
ID AAY50972 standard; Protein; 98 AA.
XX AC
XX
XX AAY50972;
XX
XX 23-MAR-2000 (first entry)
XX
XX Human FVIII antibody A2 scFv heavy chain protein DP-47 #1.
XX
XX Human; heavy chain; antibody; factor VIII; hemostatic;
XX KW hemophilia A; scFv; A2.
XX
XX Homo sapiens.
XX OS

Query Match 39.7%; Score 48; DB 21; Length 98;
Best Local Similarity 100.0%; Pred. No. 6.5e-39;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 ISGSGGTYADSVKGRFTISRDNKNTLYLQMNSLRAEDTAVYYCAK 98
|||||
Db 51 isgsggtyyadvskgrftisrdnsknltlylqmnsltraedtavyycak 98
|||||

RESULT 7
AAE12710
ID AAE12710 standard; peptide; 98 AA.
XX AC
XX
XX AAE12710;
XX
XX 04-JAN-2002 (first entry)
XX
XX DE Antibody variable heavy chain region (VH) from D-47 germ line.
XX
XX Tumour-associated antigen mucin-1; MUC-1; adenocarcinoma; VH;
XX KW variable heavy chain region; cancer; breast; ovary; bladder;
XX KW cytostatic; therapy.
XX
XX OS Unidentified.
XX
XX WO200175110-A2.
XX
XX 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US10589.
XX
XX 30-MAR-2000; 2000US-0538913.
XX
XX (DYAX-) DYAX CORP.
XX
XX Hoogenboom HRJM, Henderikx MFG;
XX
XX WPI; 2001-626437/72.
XX

```

```

XX WO9958680-A2.
XX 18-NOV-1999.
XX
XX 07-MAY-1999; 99WO-NL00285.
XX
XX 08-MAY-1998; 98EP-0201543.
XX
XX (SANO-) STICHTING SANQUIN BLOEDVOORZIENING.
XX
XX Voorberg JJ, Van Den Brink EN, Turenhout EAM;
XX
XX WPI; 2000-053102/04.
XX
XX New polynucleotide, polypeptide and antibody useful for diagnosing the
XX presence of neutralizing antibodies against factor VIII and for
XX treatment of hemophilia A patients with these antibodies -
XX
XX Example 9; Fig 11A; 61pp; English.
XX
XX This invention describes a novel polynucleotide (i) (and complements and
XX hybridizable polynucleotides) comprising a contiguous nucleotide sequence
XX coding for a human antibody with factor VIII specificity which has
XX hemostatic activity. (i) is useful as a primer or probe for detecting the
XX presence of inhibitory antibodies directed against factor VIII. The
XX polypeptides of the invention and the antibodies generated from them
XX are useful in compositions for neutralizing factor VIII inhibiting
XX antibodies in hemophilia A patients. This sequence represents a human
XX factor VIII antibody A2 specific scFv protein DP-47 which is used
XX in the method of the invention.
XX
XX Sequence 98 AA;

Query Match 39.7%; Score 48; DB 21; Length 98;
Best Local Similarity 100.0%; Pred. No. 6.5e-39;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 ISGSGGTYADSVKGRFTISRDNKNTLYLQMNSLRAEDTAVYYCAK 98
|||||
Db 51 isgsggtyyadvskgrftisrdnsknltlylqmnsltraedtavyycak 98
|||||

RESULT 7
AAE12710
ID AAE12710 standard; peptide; 98 AA.
XX AC
XX
XX AAE12710;
XX
XX 04-JAN-2002 (first entry)
XX
XX DE Antibody variable heavy chain region (VH) from D-47 germ line.
XX
XX Tumour-associated antigen mucin-1; MUC-1; adenocarcinoma; VH;
XX KW variable heavy chain region; cancer; breast; ovary; bladder;
XX KW cytostatic; therapy.
XX
XX OS Unidentified.
XX
XX WO200175110-A2.
XX
XX 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US10589.
XX
XX 30-MAR-2000; 2000US-0538913.
XX
XX (DYAX-) DYAX CORP.
XX
XX Hoogenboom HRJM, Henderikx MFG;
XX
XX WPI; 2001-626437/72.
XX

```

XX Novel isolated tumor-associated antigen mucin-1-specific binding member  
 PT for diagnosing and treating cancer, comprises mucin-1 binding domain or  
 PT its portion for binding to an epitope of the protein core of mucin-1 -  
 XX  
 XX Example 1; Page 35; 126pp; English.  
 XX  
 CC The invention relates to an isolated tumour-associated antigen mucin-1  
 CC (MUC-1)-specific binding member comprising an antigen binding domain  
 CC region having an antibody variable light (VL) or heavy (VH) region,  
 CC or a complementarity determining region (CDR) of VL or VH. MUC1-specific  
 CC binding member is useful for diagnosing cancer, preferably adenocarcinoma  
 CC The binding of MUC1-specific binding member to MUC1 is detected by a  
 CC detection method selected from enzyme-linked immunosorbent assay,  
 CC magnetic resonance imaging, scintillation counting, and X-ray film.  
 CC MUC1-specific binding member is useful for treating cancer, preferably  
 CC adenocarcinoma, in an individual, where the cancer is present in tissue  
 CC of the breast, ovary, lung, or bladder of the individual. MUC1-specific  
 CC binding member is useful for diagnosing and imaging MUC1-expressing  
 CC cancer cells and tissues, for purifying or isolating non-glycosylated,  
 CC underglycosylated or cancer-associated forms of MUC1 or MUC1 epitope-  
 CC containing molecules, and for therapeutically or prophylactically  
 CC treating cancer. The present sequence is antibody variable heavy chain  
 CC region (VH) from Dp47 germ line.  
 XX Sequence 98 AA;  
 SQ

Query Match 39.7%; Score 48; DB 22; Length 98;  
 Best Local Similarity 100.0%; Pred. No. 6.5e-39;  
 Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 51 ISGSGSTYYADSVKGRFTISRDNSKNTLYQMNSLRADETAVYYCAK 98  
 DB 51 ISGSGSTYYADSVKGRFTISRDNSKNTLYQMNSLRADETAVYYCAK 98

RESULT 8  
 AAE07017  
 ID AAE07017 standard; Protein; 116 AA.  
 XX  
 AC AAE07017;  
 XX  
 XX 16-OCT-2001 (first entry)  
 XX  
 DE Human heavy chain variable (VH) region, 034514.  
 XX  
 KW Human: humanised antibody; CC-chemokine receptor 2; CCR2; nephrotropic;  
 KW neuroprotective; immunosuppressive; human immunodeficiency virus;  
 KW HIV infection; cytostatic; vasotropic; leukocyte trafficking; allergy;  
 KW inflammatory disorder; autoimmune disorder; rheumatoid arthritis; shock;  
 KW multiple sclerosis; atherogenesis; atherosclerosis; restenosis; asthma;  
 KW anaphylaxis; malignancy; inflammation; stenosis; allograft rejection;  
 KW fibrotic disease; angioplasty; acquired immune deficiency syndrome;  
 KW AIDS; inflammatory glomerulopathy; vascular intervention; 1D9 antibody;  
 KW neointimal hyperplasia; VH; heavy chain variable region.  
 XX  
 OS Homo sapiens.

XX Key Location/Qualifiers  
 FH Region 31..35  
 FT /label= CDR1  
 FT /note= "Complementarity determining region 1"  
 FT Region 50..66  
 FT /label= CDR2  
 FT /note= "Complementarity determining region 2"  
 FT Region 99..105  
 FT /label= CDR3  
 FT /note= "Complementarity determining region 3"

WO200157226-A1.

09-AUG-2001.

XX 02-FEB-2001; 2001WO-US03537.  
 XX  
 XX 03-FEB-2000; 2000US-0497625.  
 XX  
 XX (MILL-) MILLENNIUM PHARM INC.  
 XX  
 XX Larosa GJ, Horvath C, Newman W, Jones ST, O'Brien S, O'Keefe T;  
 PI WPI; 2001-488888/53.  
 XX  
 CC Humanized immunoglobulin for treating a CC-chemokine receptor  
 CC 2-mediated disorder in a patient, comprises a binding specificity for  
 CC CCR2, and a non-human antigen binding region and human immunoglobulin  
 CC .  
 XX  
 XX Disclosure; Page 169-170; 183pp; English.  
 XX  
 CC The patent discloses a humanised antibody or its antigen-binding  
 CC fragment, having binding specificity for CC-chemokine receptor 2  
 CC (CCR2), comprising an antigen binding region of non-human origin  
 CC and at least a portion of an immunoglobulin of human origin. The  
 CC humanised antibodies are useful for inhibiting the interaction of  
 CC a cell expressing CCR2. They are useful for inhibiting or treating  
 CC HIV infection. The proteins of the invention are useful for inhibiting  
 CC leukocyte trafficking, for treating CCR2-mediated disorders such as  
 CC inflammatory disorder, autoimmune disorders such as rheumatoid  
 CC arthritis and multiple sclerosis, atherogenesis and atherosclerosis,  
 CC and for inhibiting restenosis. They are useful in therapy or diagnosis,  
 CC and in the manufacture of a medicament for treating CCR-2 mediated  
 CC disease. They are also useful for treating allergy, anaphylaxis,  
 CC malignancy, chronic and acute inflammation, histamine and IgE-  
 CC mediated allergic reaction, shock, stenosis, allograft rejection,  
 CC fibrotic disease, asthma, inflammatory glomerulopathies, acquired  
 CC immune deficiency syndrome (AIDS), restenosis associated with vascular  
 CC intervention, including angioplasty and/or stent placement in a mammal.  
 CC Humanised antibodies are also useful for inhibiting narrowing of the  
 CC lumen of a vessel in a mammal, and inhibiting neointimal hyperplasia of  
 CC a vessel in a mammal, preferably associated with vascular intervention.  
 CC The present sequence is human heavy chain variable (VH) region, 034514.  
 XX  
 XX Sequence 116 AA;

Query Match 39.7%; Score 48; DB 22; Length 116;  
 Best Local Similarity 100.0%; Pred. No. 7.5e-39;  
 Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 ISGSGSTYYADSVKGRFTISRDNSKNTLYQMNSLRADETAVYYCAK 98  
 DB 51 ISGSGSTYYADSVKGRFTISRDNSKNTLYQMNSLRADETAVYYCAK 98

RESULT 9

ID AAR66315 standard; Protein; 117 AA.

XX AAR66315;

XX 03-AUG-1995 (first entry)

XX Human immunoglobulin variable heavy chain #21.

XX Primer; PCR: amplify; human: immunoglobulin; variable; heavy chain;  
 XX cosmid; placenta; vector; pJJB81; E.coli; mammalian.

XX Homo sapiens.

XX WO9426895-A.

XX 24-NOV-1994.

XX 10-MAY-1993; 93WO-JP00603.

XX 10-MAY-1993; 93WO-JP00603.  
XX (NISH) JAPAN TOBACCO INC.  
XX Honjo T, Matsuda F;  
XX WPI; 1995-006791/01.  
XX N-PSDB; AAQ78961.  
XX DNA fragment comprising human immunoglobulin Vh genes - for the  
XX production of human immunoglobulin in mammalian hosts  
XX Claim 32; Page 60-61; 130pp; Japanese.  
XX Protein sequences (AAR66295-51) are novel human immunoglobulin heavy  
XX chain sequences encoded by novel isolated genes. The genes  
XX (AAQ78939-79002) were isolated and cloned from a series of cosmid  
XX constructs; Y202; Y103; Y21; Y6; Y24; 3-31; M84; M118 and M131, by PCR  
XX amplification using primers AAQ78917-38. The genes are subdivided into 5  
XX families of Vh genes. The fragments cover a region of 800 kb. The DNA  
XX fragments were isolated from high molecular weight DNA from human  
XX placenta. The DNA was partially digested with *NotI* restriction enzyme.  
XX The fragments were separated by gel electrophoresis and 35-45 kb fractions  
XX were collected. The fragments were ligated with *NotI*-digested cosmid  
XX vector pJB81. The ligation products were in vitro packed and infected  
XX into E.coli 490A. The fragments were then subcloned by colony  
XX hybridisation. The Vh genes and the DNA fragments encoding them are  
XX useful in producing human immunoglobulin in mammalian hosts.  
XX Sequence 117 AA;  
SQ  
Query Match 39.7%; Score 48; DB 16; Length 117;  
Best Local Similarity 100.0%; Pred. No. 7.6e-39;  
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 51 ISGSGSTYYADSVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYCAK 98  
Db 70 ISGSGSTYYADSVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYCAK 117  
RESULT 10  
AAE07022  
ID AAE07022 standard; Protein: 120 AA.  
AC AAE07022;  
XX  
XX 16-OCT-2001 (first entry)  
XX Human heavy chain variable (VH) region, VH clone 39.  
XX Human; humanised antibody; CC-chemokine receptor 2; CCR2; nephrotropic;  
XX neuroprotective; immunosuppressive; human immunodeficiency virus;  
XX HIV infection; cytostatic; vasotrophic; leukocyte trafficking; allergy;  
XX inflammatory disorder; autoimmunity; rheumatoid arthritis; shock;  
XX multiple sclerosis; atherogenesis; atherosclerosis; stenosis; asthma;  
XX anaphylaxis; malignancy; inflammation; stenosis; allograft rejection;  
XX fibrotic disease; angioplasty; acquired immune deficiency syndrome;  
XX AIDS; inflammatory glomerulopathy; vascular intervention; ID9 antibody;  
XX neonatal hyperplasia; VH; heavy chain variable region.  
XX Homo sapiens.  
XX  
XX Key Location/Qualifiers  
XX Region 31..35  
XX /label= CDR1  
XX /note= "Complementarity determining region 1"  
XX Region 50..66  
XX /label= CDR2  
XX /note= "Complementarity determining region 2"  
XX Region 99..109  
XX /label= CDR3

FT /note= "Complementarity determining region 3"  
XX WO200157226-A1.  
XX 09-AUG-2001.  
XX 02-FEB-2001; 2001WO-US03537.  
XX 03-FEB-2000; 2000US-0497625.  
XX (MILL-) MILLENNIUM PHARM INC.  
XX Larosa GJ, Horvath C, Newman W, Jones ST, O'Brien S, O'Keefe T;  
XX WPI; 2001-488888/53.  
XX Humanized immunoglobulin for treating a CC-chemokine receptor  
XX 2-mediated disorder in a patient, comprises a binding specificity for  
XX CCR2, and a non-human antigen binding region and human immunoglobulin  
XX -  
XX Disclosure; Page 171-172; 183pp; English.  
XX The patent discloses a humanised antibody or its antigen-binding  
XX fragment, having binding specificity for CC-chemokine receptor 2  
XX (CCR2), comprising an antigen binding region of non-human origin  
XX and at least a portion of an immunoglobulin of human origin. The  
XX humanised antibodies are useful for inhibiting the interaction of  
XX a cell expressing CCR2. They are useful for inhibiting or treating  
XX HIV infection. The proteins of the invention are useful for inhibiting  
XX leukocyte trafficking, for treating CCR2-mediated disorders such as  
XX inflammatory disorder, autoimmune disorders such as rheumatoid  
XX arthritis and multiple sclerosis, atherogenesis and atherosclerosis,  
XX and for inhibiting stenosis. They are useful in therapy or diagnosis,  
XX and in the manufacture of a medicament for treating CCR2 mediated  
XX disease. They are also useful for treating allergy, anaphylaxis,  
XX malignancy, chronic and acute inflammation, histamine and IgE-  
XX mediated allergic reaction, shock, stenosis, allograft rejection,  
XX fibrotic disease, asthma, inflammatory glomerulopathies, acquired  
XX immune deficiency syndrome (AIDS), restenosis associated with vascular  
XX intervention, including angioplasty and/or stent placement in a mammal.  
XX Humanised antibodies are also useful for inhibiting narrowing of the  
XX lumen of a vessel in a mammal, and inhibiting neointimal hyperplasia of  
XX a vessel in a mammal, preferably associated with vascular intervention.  
XX The present sequence is human heavy chain variable (VH) region, VH  
XX clone 39.  
XX Sequence 120 AA;  
SQ  
Query Match 39.7%; Score 48; DB 22; Length 120;  
Best Local Similarity 100.0%; Pred. No. 7.7e-39;  
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 51 ISGSGSTYYADSVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYCAK 98  
Db 51 ISGSGSTYYADSVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYCAK 98  
RESULT 11  
AAE07029  
ID AAE07029 standard; Protein: 121 AA.  
XX  
XX AAE07029;  
XX  
XX 16-OCT-2001 (first entry)  
XX Human heavy chain variable (VH) region, VH clone 41.  
XX Human; humanised antibody; CC-chemokine receptor 2; CCR2; nephrotropic;  
XX neuroprotective; immunosuppressive; human immunodeficiency virus;  
XX HIV infection; cytostatic; vasotrophic; leukocyte trafficking; allergy;  
XX inflammatory disorder; autoimmunity; rheumatoid arthritis; shock;  
XX



KW multiple sclerosis; atherogenesis; atherosclerosis; restenosis; asthma;  
KW anaphylaxis; malignancy; inflammation; stenosis; allograft rejection;  
KW fibrotic disease; angioplasty; acquired immune deficiency syndrome;  
KW AIDS; inflammatory glomerulopathy; vascular intervention; ID9 antibody;  
KW neonatal hyperplasia; VH; heavy chain variable region.

OS Homo sapiens.

XX Key Location/Qualifiers  
FH Region 31..35  
FT /label= CDR1  
FT /note= "Complementarity determining region 1"  
FT 50..66  
FT /label= CDR2  
FT /note= "Complementarity determining region 2"  
FT 99..110  
FT /label= CDR3  
FT /note= "Complementarity determining region 3"

XX WO200157226-A1.

PN 09-AUG-2001.

PD 02-FEB-2001; 2001WO-US03537.

PF 03-FEB-2000; 2000US-0497625.

PR (MILL-) MILLENNIUM PHARM INC.

PA Larosa GJ, Horvath C, Newman W, Jones ST, O'Brien S, O'Keefe T;

PI WPI; 2001-488888/53.

XX Humanized immunoglobulin for treating a CC-chemokine receptor  
PT 2-mediated disorder in a patient, comprises a binding specificity for  
PT CCR2, and a non-human antigen binding region and human immunoglobulin  
PT -

PS Disclosure; Page 174-175; 183pp; English.

XX The patent discloses a humanised antibody or its antigen-binding  
CC fragment, having binding specificity for CC-chemokine receptor 2  
CC (CCR2), comprising an antigen binding region of non-human origin  
CC and at least a portion of an immunoglobulin of human origin. The  
CC humanised antibodies are useful for inhibiting the interaction of  
CC a cell expressing CCR2. They are useful for inhibiting or treating  
CC HIV infection. The proteins of the invention are useful for inhibiting  
CC leukocyte trafficking, for treating CCR2-mediated disorders such as  
CC inflammatory disorder, autoimmune disorders such as rheumatoid  
CC arthritis and multiple sclerosis, atherogenesis and atherosclerosis,  
CC and for inhibiting restenosis. They are useful in therapy or diagnosis,  
CC and in the manufacture of a medicament for treating CCR-2 mediated  
CC disease. They are also useful for treating allergy, anaphylaxis,  
CC malignancy, chronic and acute inflammation, histamine and IgE-  
CC mediated allergic reaction, shock, stenosis, allograft rejection,  
CC fibrotic disease, asthma, inflammatory glomerulopathies, acquired  
CC immune deficiency syndrome (AIDS), restenosis associated with vascular  
CC intervention, including angioplasty and/or stent placement in a mam-  
CC mal. Humanised antibodies are also useful for inhibiting narrowing of the  
CC lumen of a vessel in a mammal, and inhibiting neointimal hyperplasia of  
CC a vessel in a mammal, preferably associated with vascular intervention.  
CC The present sequence is human heavy chain variable (VH) region,  
CC VH clone 41.

XX Sequence 121 AA;

Query Match 39.7%; Score 48; DB 22; Length 121;  
Best Local Similarity 100.0%; Pred. No. 7.8e-39;  
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 ISGSGSTYYADSVKGRFTISRDNKNTLYLQMSLRADTAVYYCAK 98  
|||||

Db 51 isgsgstyyadvskgrftisrdnsknntlylqmslraedtavyycaak 98

RESULT 12

AAU02589

ID AAU02589 standard; Protein; 122 AA.

XX AC AAU02589;

XX DT 29-AUG-2001 (first entry)

XX DE Anti-adipocyte monoclonal antibody heavy chain, PAT 75.

XX KW Antibody; adipocyte; heavy chain; light chain; obesity; fat;

XX KW heart disease; complementarity determining region; CDR.

XX OS Homo sapiens.

PN WO200127279-A1.

XX PD 19-APR-2001.

XX PF 11-OCT-2000; 2000WO-GB03900.

XX PR 12-OCT-1999; 99US-0158812.

XX PA (CAMP-) CAMBRIDGE ANTIBODY TECHNOLOGY.

XX PI Edwards BM, Main SH, Vaughan TJ;

XX DR WPI; 2001-282031/29.

XX DR N-PSDB; AAS03489.

XX Panel of specific binding members of antibody molecules which bind to  
PT whole adipocytes is used in the treatment of obesity and obesity  
PT related diseases -

XX Claim 1; Page 148; 182pp; English.

XX AAU02501-AAU02635, and AAU02641-AAU02748 represent the amino acid  
CC sequences of anti-adipocyte monoclonal antibody heavy chain, light  
CC chain, and heavy chain complementarity determining regions (CDR) of the  
CC invention. The antibodies can be used in the treatment of obesity and  
CC obesity related diseases. The antibodies can be used to deliver drugs or  
CC pro-drugs directly to the fat mass of an obese patient or the antibody  
CC can be used as a therapeutic itself. Antibodies binding specifically to  
CC adipocytes can be used to activate the immune system to destroy the cells  
CC by complement mediated lysis. The antibodies may be labeled with a  
CC detectable label such as radiolabel, fluorescent or chemical group and  
CC used in methods of diagnosis in human subjects e.g. to determine the  
CC presence of adipocyte antigen on the surface of an adipocyte to detect or  
CC determine the presence or level of adipocytes in a cell or tissue sample.  
CC The antibodies can be used as an alternative means of treatment for obese  
CC patients other than undergoing surgery to remove excess fat. Antibodies  
CC for different types of fat deposits can also be produced e.g. intra-  
CC abdominal fat associated with heart disease.

XX Sequence 122 AA;

Query Match 39.7%; Score 48; DB 22; Length 122;  
Best Local Similarity 100.0%; Pred. No. 7.8e-35;  
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 ISGSGSTYYADSVKGRFTISRDNKNTLYLQMSLRADTAVYYCAK 98  
|||||

Db 51 isgsgstyyadvskgrftisrdnsknntlylqmslraedtavyycaak 98  
|||||

RESULT 13

AAE07013

ID AAE07013 standard; Protein; 125 AA.

XX

AC AAE07013;  
XX  
XX 16-OCT-2001 (first entry)  
XX Human heavy chain variable (VH) region, 038064.  
XX  
XX Human; humanised antibody; CC-chemokine receptor 2; CCR2; nephrotropic;  
KW neuroprotective; immunosuppressive; human immunodeficiency virus;  
KW HIV infection; cytostatic; vasotropic; leukocyte trafficking; allergy;  
KW inflammatory disorder; autoimmune disorder; rheumatoid arthritis; shock;  
KW multiple sclerosis; atherogenesis; atherosclerosis; restenosis; asthma;  
KW anaphylaxis; malignancy; inflammation; stenosis; allograft rejection;  
KW fibrotic disease; angiotomy; acquired immune deficiency syndrome;  
KW AIDS; inflammatory glomerulopathy; vascular intervention; ID9 antibody;  
KW neointimal hyperplasia; VH; heavy chain variable region.  
XX  
XX Homo sapiens.  
XX  
XX Key Location/Qualifiers  
XX Region 31..35  
XX /label= CDR1  
XX /note= "Complementarity determining region 1"  
XX Region 50..66  
XX /label= CDR2  
XX /note= "Complementarity determining region 2"  
XX Region 99..114  
XX /label= CDR3  
XX /note= "Complementarity determining region 3"  
XX WO200157226-A1.  
XX  
XX 09-AUG-2001.  
XX  
XX 02-FEB-2001; 2001WO-US03537.  
XX  
XX 03-FEB-2000; 2000US-0497625.  
XX (MILL-) MILLENNIUM PHARM INC.  
XX  
XX Larosa GJ, Horvath C, Newman W, Jones ST, O'Brien S, O'Keefe T;  
XX WPI; 2001-488888/53.  
XX  
XX Humanized immunoglobulin for treating a CC-chemokine receptor  
XX 2-mediated disorder in a patient, comprises a binding specificity for  
XX CCR2, and a non-human antigen binding region and human immunoglobulin  
XX  
XX Disclosure; Page 168; 183pp; English.  
XX  
XX The patent discloses a humanised antibody or its antigen-binding  
XX fragment, having binding specificity for CC-chemokine receptor 2  
XX (CCR2), comprising an antigen binding region of non-human origin  
XX and at least a portion of an immunoglobulin of human origin. The  
XX humanised antibodies are useful for inhibiting the interaction of  
XX a cell expressing CCR2. They are useful for inhibiting or treating  
XX HIV infection. The proteins of the invention are useful for inhibiting  
XX leukocyte trafficking, for treating CCR2-mediated disorders such as  
XX inflammatory disorder, autoimmune disorders such as rheumatoid  
XX arthritis and multiple sclerosis, atherogenesis and atherosclerosis,  
XX and for inhibiting restenosis. They are useful in therapy or diagnosis,  
XX and in the manufacture of a medicament for treating CCR-2 mediated  
XX disease. They are also useful for treating allergy, anaphylaxis,  
XX mediated allergic reaction, shock, stenosis, allograft rejection,  
XX fibrotic disease, asthma, inflammatory glomerulopathies, acquired  
XX immune deficiency syndrome (AIDS), restenosis associated with vascular  
XX intervention, including angiotomy and/or stent placement in a mammal.  
XX Humanised antibodies are also useful for inhibiting narrowing of the  
XX lumen of a vessel in a mammal, and inhibiting neointimal hyperplasia of  
XX a vessel in a mammal, preferably associated with vascular intervention.  
XX The present sequence is human heavy chain variable (VH) region, 038064.  
XX

SQ Sequence 125 AA;  
Query Match 39.7%; Score 48; DB 22; E-Val 125;  
Best Local Similarity 100.0%; Pred. No. 8e-39;  
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 51 ISGSGSTYYADSVKGRFTISRDNSKNTLYLOMNSLRADETAVYVAK 98  
DB 51 ISGSGSTYYADSVKGRFTISRDNSKNTLYLOMNSLRADETAVYVAK 98  
RESULT 14  
AAE07021  
ID AAE07021 standard; Protein; 125 AA.  
XX  
XX AAE07021;  
XX  
XX 16-OCT-2001 (first entry)  
XX Human heavy chain variable (VH) region, 4G12.  
XX  
XX Human; humanised antibody; CC-chemokine receptor 2; CCR2; nephrotropic;  
KW neuroprotective; immunosuppressive; human immunodeficiency virus;  
KW HIV infection; cytostatic; vasotropic; leukocyte trafficking; allergy;  
KW inflammatory disorder; autoimmune disorder; rheumatoid arthritis; shock;  
KW multiple sclerosis; atherogenesis; atherosclerosis; restenosis; asthma;  
KW anaphylaxis; malignancy; inflammation; stenosis; allograft rejection;  
KW fibrotic disease; angiotomy; acquired immune deficiency syndrome;  
KW AIDS; inflammatory glomerulopathy; vascular intervention; ID9 antibody;  
KW neointimal hyperplasia; VH; heavy chain variable region.  
XX  
XX Homo sapiens.  
XX  
XX Key Location/Qualifiers  
XX Region 31..35  
XX /label= CDR1  
XX /note= "Complementarity determining region 1"  
XX Region 50..66  
XX /label= CDR2  
XX /note= "Complementarity determining region 2"  
XX Region 99..114  
XX /label= CDR3  
XX /note= "Complementarity determining region 3"  
XX WO200157226-A1.  
XX  
XX 09-AUG-2001.  
XX  
XX 02-FEB-2001; 2001WO-US03537.  
XX  
XX 03-FEB-2000; 2000US-0497625.  
XX (MILL-) MILLENNIUM PHARM INC.  
XX  
XX Larosa GJ, Horvath C, Newman W, Jones ST, O'Brien S, O'Keefe T;  
XX WPI; 2001-488888/53.  
XX  
XX Humanized immunoglobulin for treating a CC-chemokine receptor  
XX 2-mediated disorder in a patient, comprises a binding specificity for  
XX CCR2, and a non-human antigen binding region and human immunoglobulin  
XX  
XX Disclosure; Page 171; 183pp; English.  
XX  
XX The patent discloses a humanised antibody or its antigen-binding  
XX fragment, having binding specificity for CC-chemokine receptor 2  
XX (CCR2), comprising an antigen binding region of non-human origin  
XX and at least a portion of an immunoglobulin of human origin. The  
XX humanised antibodies are useful for inhibiting the interaction of  
XX a cell expressing CCR2. They are useful for inhibiting or treating  
XX HIV infection. The proteins of the invention are useful for inhibiting  
XX leukocyte trafficking, for treating CCR2-mediated disorders such as  
XX inflammatory disorder, autoimmune disorders such as rheumatoid  
XX arthritis and multiple sclerosis, atherogenesis and atherosclerosis,  
XX and for inhibiting restenosis. They are useful in therapy or diagnosis,  
XX and in the manufacture of a medicament for treating CCR-2 mediated  
XX disease. They are also useful for treating allergy, anaphylaxis,  
XX mediated allergic reaction, shock, stenosis, allograft rejection,  
XX fibrotic disease, asthma, inflammatory glomerulopathies, acquired  
XX immune deficiency syndrome (AIDS), restenosis associated with vascular  
XX intervention, including angiotomy and/or stent placement in a mammal.  
XX Humanised antibodies are also useful for inhibiting narrowing of the  
XX lumen of a vessel in a mammal, and inhibiting neointimal hyperplasia of  
XX a vessel in a mammal, preferably associated with vascular intervention.  
XX The present sequence is human heavy chain variable (VH) region, 038064.  
XX

CC leukocyte trafficking, for treating CCR2-mediated disorders such as  
 CC inflammatory disorder, autoimmune disorders such as rheumatoid  
 CC arthritis and multiple sclerosis, atherogenesis and atherosclerosis,  
 CC and for inhibiting restenosis. They are useful in therapy or diagnosis,  
 CC and in the manufacture of a medicament for treating CCR-2 mediated  
 CC disease. They are also useful for treating allergy, anaphylaxis,  
 CC malignancy, chronic and acute inflammation, histamine and IgE-  
 CC mediated allergic reaction, shock, stenosis, allograft rejection,  
 CC fibrotic disease, asthma, inflammatory glomerulopathies, acquired  
 CC immune deficiency syndrome (AIDS), restenosis associated with vascular  
 CC intervention, including angioplasty and/or stent placement in a mammal.  
 CC Humanised antibodies are also useful for inhibiting narrowing of the  
 CC lumen of a vessel in a mammal, and inhibiting neointimal hyperplasia of  
 CC a vessel in a mammal, preferably associated with vascular intervention.  
 CC The present sequence is human heavy chain variable (VH) region, 4G12.

XX Sequence 125 AA;

Query Match 39.7%; Score 48; DB 22; Length 125;  
 Best Local Similarity 100.0%; Pred. No. 8e-39;  
 Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 ISGSGSTYYADSVKGRFTISRDNKNTLYLQMNSLRAEDTAVYYCAK 98  
 |||||  
 Db 51 ISGSGSTYYADSVKGRFTISRDNKNTLYLQMNSLRAEDTAVYYCAK 98

#### RESULT 15

AAW06242  
 ID AAW06242 standard; Protein; 128 AA.

XX AC AAW06242;

XX DT 06-AUG-1997 (first entry)

XX DE Heavy chain variable CDR3 region of anti-lymphoma antibody.

XX KW Ig; immunoglobulin; antibody; B cell; heavy chain variable region;  
 KW CDR; complementarity determining region; malignant; cancer; tumour;  
 KW neoplasia; cell mediated immune enhancement; lymphoma.

XX OS Homo sapiens.

XX PN WO9636714-A1.

XX PD 21-NOV-1996.

XX PF 14-MAY-1996; 96WO-JP01258.

XX PR 18-MAY-1995; 95JP-0119589.

XX PA (TAKI ) TAKARA SHUZO CO LTD.

XX PI Abe A, Emi N;

XX DR WPI; 1997-012089/01.

XX DR N-PSDB; AAT43529.

XX PT Nucleic acid expressing CDR III region of immunoglobulin VH region -  
 PT also vector containing it, for treatment of B-cell malignant tumours

XX PS Claim 3; Page 26-27; 40pp; English.

XX CC AAW06242 represents the heavy chain variable CDR 3 region of an antibody  
 CC specific to human B-cell malignant tumour cells. The nucleic acid  
 CC molecule may be used in the treatment of human B-cell malignant  
 CC tumours (lymphomas), e.g. by cell-mediated immune enhancement via  
 CC administration of the nucleic acid molecule to patients (preferably  
 CC in a vector such as a retrovirus).

XX SQ Sequence 128 AA;

Query Match 39.7%; Score 48; DB 18; Length 128;  
 Best Local Similarity 100.0%; Pred. No. 8.2e-35;  
 Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 ISGSGSTYYADSVKGRFTISRDNKNTLYLQMNSLRAEDTAVYYCAK 98  
 |||||  
 Db 50 ISGSGSTYYADSVKGRFTISRDNKNTLYLQMNSLRAEDTAVYYCAK 97

Search completed: October 4, 2002, 10:29:42  
 Job time: 234 sec

---

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 4, 2002, 10:19:12 ; Search time 20.58 seconds  
(without alignments)  
143.610 Million cell updates/sec

Title: US-09-822-698A-3  
Perfect score: 644  
Sequence: 1 QVQLVQSGSLVPGSLRL.....GGVWDPIYWGQGLTVSS 121

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 2442594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA:  
1: /cgn2\_6/ptodata/2/iaa/5A-COMB.pep.\*  
2: /cgn2\_6/ptodata/2/iaa/5A-COMB.pep.\*  
3: /cgn2\_6/ptodata/2/iaa/6A-COMB.pep.\*  
4: /cgn2\_6/ptodata/2/iaa/6A-COMB.pep.\*  
5: /cgn2\_6/ptodata/2/iaa/6B-COMB.pep.\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	543.5	84.4	120	4	US-09-025-769B-38
2	543.5	84.4	120	4	US-09-025-769B-63
3	543.5	84.4	281	4	US-09-025-769B-178
4	537.5	83.5	131	4	US-08-983-607-28
5	535	83.1	117	4	US-08-983-607-46
6	527	81.8	120	2	US-08-428-197-20
7	527	81.8	120	5	PCT-US93-10555-20
8	527	81.8	125	2	US-08-428-197-1
9	527	81.8	125	5	PCT-US93-10555-1
10	525	81.5	120	2	US-08-428-197-22
11	525	81.5	120	5	PCT-US93-10555-22
12	524.5	81.4	116	4	US-08-983-607-36
13	524.5	81.4	122	2	US-07-934-373C-21
14	524.5	81.4	122	3	US-08-437-642B-21
15	524.5	81.4	122	5	PCT-US93-07832-21
16	524	81.4	120	2	US-08-428-197-24
17	524	81.4	120	2	US-08-428-197-26
18	524	81.4	120	2	US-08-428-197-28
19	524	81.4	120	5	PCT-US93-10555-24
20	524	81.4	120	5	PCT-US93-10555-26
21	524	81.4	120	5	PCT-US93-10555-28
22	523	81.2	113	3	US-08-974-899-6
23	520.5	80.8	140	4	US-08-983-607-32
24	520	80.7	117	4	US-09-025-769B-24
25	517	80.3	125	1	US-08-478-039-99
26	517	80.3	125	1	US-08-476-349A-99
27	514	79.8	120	2	US-08-428-197-40

28	514	79.8	120	5	PCT-US93-10555-40
29	512.5	79.6	283	4	US-09-420-592A-6
30	511	79.3	119	1	US-07-988-925-11
31	511	79.3	119	2	US-08-362-780-11
32	511	79.3	120	2	US-08-428-197-30
33	511	79.3	120	2	US-08-428-197-32
34	511	79.3	120	5	PCT-US93-10555-30
35	511	79.3	120	5	PCT-US93-10555-32
36	510.5	79.3	123	5	PCT-US93-10555-38
37	510.5	79.3	123	5	PCT-US93-10555-38
38	508	78.9	120	2	US-08-428-197-34
39	508	78.9	120	5	PCT-US93-10555-34
40	508	78.9	123	2	US-08-565-202-30
41	505.5	78.5	116	5	PCT-US93-10555-2
42	505.5	78.5	118	2	US-08-958-201-2
43	505.5	78.5	118	2	US-08-958-201-4
44	505.5	78.5	118	2	US-08-958-201-4
45	504.5	78.3	118	2	US-08-652-816A-12

## ALIGNMENTS

RESULT 1  
US-09-025-769B-38  
; Sequence 38, Application US/09025769B  
; Patent No. 6300064  
; GENERAL INFORMATION:  
; APPLICANT: Knappik, Achim  
; APPLICANT: Pack, Peter  
; APPLICANT: Ilag, Vic  
; APPLICANT: Ge, Liming  
; APPLICANT: Moroney, Simon  
; APPLICANT: Plueckhuhn, Andreas  
; TITLE OF INVENTION: Protein/(Poly)peptide libraries  
; NUMBER OF SEQUENCES: 373  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: James F. Haley, Jr., Esq, c/o Fish & Neave  
; STREET: 1251 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10021  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.3.0 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/025,769B  
; FILING DATE: 18-FEB-1998  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: EP 95 11 3021.0  
; FILING DATE: 18-AUG-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: James F. Haley, Jr., Esq.  
; REGISTRATION NUMBER: 27,794  
; REFERENCE/DOCKET NUMBER: MORPHO/5  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212)596-9000  
; TELEFAX: (212)596-9090  
; INFORMATION FOR SEQ ID NO: 38:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 120 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-09-025-769B-38  
  
Query Match 84.4%; Score 543.5; DB 4; Length 120;  
Best Local Similarity 87.6%; Pred. No. 2.3e-46;

Matches 106; Conservative 5; Mismatches 9; Indels 1; Gaps 1;

QY 1 QVQLVQSGGGLVPGGSLRLSCAASGFTFRSNAMGWVRQAPGKGLWVSGISGGSTYY 60  
Db 1 EVQLVESGGGLVPGGSLRLSCAASGFTFRSNAMGWVRQAPGKGLWVSAISGGSTYY 60  
QY 61 ADSVKGRFTISRDNKNTLYLQMSLRAEDTAVYYCAKHTGGVWDPIDYWGQGTFLVTS 120  
Db 61 ADSVKGRFTISRDNKNTLYLQMSLRAEDTAVYYCARWGDDGFY-ANDYWGQGTFLVTS 119  
QY 121 S 121  
Db 120 S 120

## RESULT 2

US-09-025-769B-63  
; Sequence 63, Application US/09025769B  
; Patent No. 6300064

## GENERAL INFORMATION:

; APPLICANT: Knappik, Achim  
; APPLICANT: Pack, Peter  
; APPLICANT: Ilag, Vic  
; APPLICANT: Ge, Liming  
; APPLICANT: Moroney, Simon  
; APPLICANT: Plueckthun, Andreas  
; TITLE OF INVENTION: Protein/(Poly)peptide libraries  
; NUMBER OF SEQUENCES: 373  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: James F. Haley, Jr., Esq. c/o Fish & Neave  
; STREET: 1251 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10021

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/025,769B  
; FILING DATE: 18-FEB-1998  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: EP 95 11 3021.0  
; FILING DATE: 18-AUG-1995

ATTORNEY/AGENT INFORMATION:  
; NAME: James F. Haley, Jr., Esq.  
; REGISTRATION NUMBER: 27,794  
; REFERENCE/DOCKET NUMBER: MORPHO/5  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212)596-9000  
; TELEFAX: (212)596-9090  
; INFORMATION FOR SEQ ID NO: 63:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 120 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-025-769B-63

## Query Match

Best Local Similarity 84.4%; Score 543.5; DB 4; Length 120;

Matches 106; Conservative 5; Mismatches 9; Indels 1; Gaps 1;

QY 1 QVQLVQSGGGLVPGGSLRLSCAASGFTFRSNAMGWVRQAPGKGLWVSGISGGSTYY 60  
Db 1 EVQLVESGGGLVPGGSLRLSCAASGFTFRSNAMGWVRQAPGKGLWVSAISGGSTYY 60  
QY 61 ADSVKGRFTISRDNKNTLYLQMSLRAEDTAVYYCAKHTGGVWDPIDYWGQGTFLVTS 120  
Db 61 ADSVKGRFTISRDNKNTLYLQMSLRAEDTAVYYCARWGDDGFY-ANDYWGQGTFLVTS 119

QY 121 S 121  
Db 120 S 120

## RESULT 3

US-09-025-769B-178  
; Sequence 178, Application US/09025769B  
; Patent No. 6300064

## GENERAL INFORMATION:

; APPLICANT: Knappik, Achim  
; APPLICANT: Pack, Peter  
; APPLICANT: Ilag, Vic  
; APPLICANT: Ge, Liming  
; APPLICANT: Moroney, Simon  
; APPLICANT: Plueckthun, Andreas  
; TITLE OF INVENTION: Protein/(Poly)peptide libraries  
; NUMBER OF SEQUENCES: 373  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: James F. Haley, Jr., Esq. c/o Fish & Neave  
; STREET: 1251 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10021

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/025,769B  
; FILING DATE: 18-FEB-1998  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: EP 95 11 3021.0  
; FILING DATE: 18-AUG-1995

ATTORNEY/AGENT INFORMATION:  
; NAME: James F. Haley, Jr., Esq.  
; REGISTRATION NUMBER: 27,794  
; REFERENCE/DOCKET NUMBER: MORPHO/5  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212)596-9000  
; TELEFAX: (212)596-9090  
; INFORMATION FOR SEQ ID NO: 178:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 281 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-025-769B-178

## Query Match

Best Local Similarity 84.4%; Score 543.5; DB 4; Length 281;

Matches 106; Conservative 5; Mismatches 9; Indels 1; Gaps 1;

QY 1 QVQLVQSGGGLVPGGSLRLSCAASGFTFRSNAMGWVRQAPGKGLWVSGISGGSTYY 60  
Db 26 EVQLVESGGGLVPGGSLRLSCAASGFTFRSNAMGWVRQAPGKGLWVSAISGGSTYY 85  
QY 61 ADSVKGRFTISRDNKNTLYLQMSLRAEDTAVYYCAKHTGGVWDPIDYWGQGTFLVTS 120  
Db 86 ADSVKGRFTISRDNKNTLYLQMSLRAEDTAVYYCARWGDDGFY-ANDYWGQGTFLVTS 144  
QY 121 S 121  
Db 145 S 145

## RESULT 4

US-08-983-607-28  
; Sequence 28, Application US/08983607

PATENT NO. 6140470  
GENERAL INFORMATION:  
APPLICANT: Alan Garen  
TITLE OF INVENTION: Human Anti-Tumor Monoclonal Anti-  
TITLE OF INVENTION: bodies  
NUMBER OF SEQUENCES: 51  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Department of Molecular Biophysics  
ADDRESSEE: and Biochemistry, Yale University  
STREET: 266 Whitney Avenue  
CITY: New Haven  
STATE: Connecticut  
COUNTRY: United States of America  
ZIP: 06520-8114  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" 1.44 Mb diskette  
OPERATING SYSTEM: MS DOS  
SOFTWARE: Word Processing  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/983,607  
FILING DATE: April 27, 1998  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/IB96/01032  
FILING DATE: June 28, 1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Mary M. Krinsky  
REGISTRATION NUMBER: 32423  
REFERENCE/DOCKET NUMBER: OCR-679  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 203-773-9544  
TELEFAX: 203-773-1183  
INFORMATION FOR SEQ ID NO: 28:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 131 residues  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE:  
DESCRIPTION: polypeptide  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens (melanoma patient immu-  
INDIVIDUAL ISOLATE: peripheral blood lymphocytes  
IMMEDIATE SOURCE:  
LIBRARY: DM414 scfv antibodies obtained from  
LIBRARY: fuses fusion phase construct  
CLONE: V13  
FEATURE:  
NAME/KEY: heavy chain  
US-08-983-607-28

Query Match 83.5%; Score 537.5; DB 4; Length 131;  
Best Local Similarity 89.3%; Pred. No. 9.7e-46;  
Matches 108; Conservative 1; Mismatches 7; Indels 5; Gaps 2;

QY 1 QVQLVQSGGGLVQPQGGSLRLSCAASGFTFRSNAMGWVROAPGKLEWVSGISGSGSTYY 60  
Db :|||||  
QY 61 ADSVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYCAKHPTGGGWDPIDYWGGGLTVTVS 120  
Db :|||||  
QY 121 S 121  
Db 116 S 116

RESULT 5  
US-08-983-607-46  
Sequence 46, Application US/08983607  
Patent No. 6140470  
GENERAL INFORMATION:  
APPLICANT: Alan Garen  
TITLE OF INVENTION: Human Anti-Tumor Monoclonal Anti-  
TITLE OF INVENTION: bodies  
NUMBER OF SEQUENCES: 51  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Department of Molecular Biophysics  
ADDRESSEE: and Biochemistry, Yale University  
STREET: 266 Whitney Avenue  
CITY: New Haven  
STATE: Connecticut  
COUNTRY: United States of America  
ZIP: 06520-8114  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" 1.44 Mb diskette  
OPERATING SYSTEM: MS DOS  
SOFTWARE: Word Processing  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/983,607  
FILING DATE: April 27, 1998  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/IB96/01032  
FILING DATE: June 28, 1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Mary M. Krinsky  
REGISTRATION NUMBER: 32423  
REFERENCE/DOCKET NUMBER: OCR-679  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 203-773-9544  
TELEFAX: 203-773-1183  
INFORMATION FOR SEQ ID NO: 46:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 117 residues  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE:  
DESCRIPTION: polypeptide  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens (melanoma patient  
INDIVIDUAL ISOLATE: peripheral blood lympho-  
INDIVIDUAL ISOLATE: cytes  
IMMEDIATE SOURCE:  
LIBRARY: VH antibodies obtained from fuses  
LIBRARY: fusion phase construct  
CLONE: C55  
FEATURE:  
NAME/KEY: heavy chain  
US-08-983-607-46

Query Match 83.18%; Score 535; DB 4; Length 117;  
Best Local Similarity 88.4%; Pred. No. 1.5e-45;  
Matches 107; Conservative 2; Mismatches 8; Indels 4; Gaps 2;

QY 1 QVQLVQSGGGLVQPQGGSLRLSCAASGFTFRSNAMGWVROAPGKLEWVSGISGSGSTYY 60  
Db :|||||  
QY 61 ADSVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYCAKHPTGGGWDPIDYWGGGLTVTVS 120  
Db :|||||  
QY 121 S 121





APPLICATION NUMBER: PCT/US93/10555  
FILING DATE: 29-OCT-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Howells, Stacy L.  
REGISTRATION NUMBER: 34,842  
REFERENCE/DOCKET NUMBER: FD-2630  
TELEPHONE: (619) 455-5100  
TELEFAX: (619) 455-5110  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 125 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
IMMEDIATE SOURCE:  
CLONE: 18/2  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: 1..125  
PCT-US93-10555-1

Query Match 81.8%; Score 527; DB 2; Length 125;  
Best Local Similarity 83.2%; Pred. No. 9.8e-45;  
Matches 104; Conservative 5; Mismatches 12; Indels 4; Gaps 1;  
US-08-428-197-1  
QY 1 QVOLVQSGGLVQPGGSLRLSCAASGFTFRSNAMGWVRQAPGKLEWVSGISGGSGSTYY 60  
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 1 EVOLLESGGLVQPGGSLRLSCAASGFTFSYAMSWVRQAPGKLEWVSAISGGSGSTYY 60  
QY 61 ADVKGRFTISRDNSKNTLYLQMSLRADETAVYYCAK----HTGGGVWDPTDYWGQGTLL 116  
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 61 ADVKGRFTISRDNSKNTLYLQMSLRADETAVYYCTKGQVLYYSGSGSIHWFDPWGQGTLL 120  
QY 117 VTSS 121  
:|||||  
Db 121 VTSS 125

RESULT 9  
PCT-US93-10555-1  
Sequence 1, Application PC/TUS9310555  
GENERAL INFORMATION:  
APPLICANT: SILVERMAN, GREGG J.  
TITLE OF INVENTION: METHOD FOR STIMULATING PRODUCTION OF  
TITLE OF INVENTION: VARIABLE REGION GENE FAMILY RESTRICTED ANTIBODIES THROUGH  
TITLE OF INVENTION: VACCINATION WITH A B-CELL SUPERANTIGEN AND CONJUGATES  
TITLE OF INVENTION: THEREOF  
NUMBER OF SEQUENCES: 51  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Spensley Horn Jubas & Lubitz  
STREET: 1880 Century Park East - Suite 500  
CITY: Los Angeles  
STATE: California  
COUNTRY: USA  
ZIP: 90067  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/10555  
FILING DATE: 29-OCT-1993  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Howells, Stacy L.  
REGISTRATION NUMBER: 34,842  
REFERENCE/DOCKET NUMBER: FD-2630  
TELEPHONE: (619) 455-5100

TELEFAX: (619) 455-5110  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 125 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
IMMEDIATE SOURCE:  
CLONE: 18/2  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: 1..125  
PCT-US93-10555-1

Query Match 81.8%; Score 527; DB 5; Length 125;  
Best Local Similarity 83.2%; Pred. No. 9.8e-45;  
Matches 104; Conservative 5; Mismatches 12; Indels 4; Gaps 1;  
QY 1 QVOLVQSGGLVQPGGSLRLSCAASGFTFRSNAMGWVRQAPGKLEWVSGISGGSGSTYY 60  
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 1 EVOLLESGGLVQPGGSLRLSCAASGFTFSYAMSWVRQAPGKLEWVSAISGGSGSTYY 60  
QY 61 ADVKGRFTISRDNSKNTLYLQMSLRADETAVYYCAK----HTGGGVWDPTDYWGQGTLL 116  
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 61 ADVKGRFTISRDNSKNTLYLQMSLRADETAVYYCTKGQVLYYSGSGSIHWFDPWGQGTLL 120  
QY 117 VTSS 121  
:|||||  
Db 121 VTSS 125

RESULT 10  
US-08-428-197-22  
Sequence 22, Application US/08428197  
Patent No. 5891438  
GENERAL INFORMATION:  
APPLICANT: SILVERMAN, GREGG J.  
TITLE OF INVENTION: METHOD FOR STIMULATING PRODUCTION OF  
TITLE OF INVENTION: VARIABLE REGION GENE FAMILY RESTRICTED ANTIBODIES THROUGH  
TITLE OF INVENTION: VACCINATION WITH A B-CELL SUPERANTIGEN AND CONJUGATES  
TITLE OF INVENTION: THEREOF  
NUMBER OF SEQUENCES: 51  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Spensley Horn Jubas & Lubitz  
STREET: 1880 Century Park East - Suite 500  
CITY: Los Angeles  
STATE: California  
COUNTRY: USA  
ZIP: 90067  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/428,197  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA: PCT/US93/10555  
APPLICATION NUMBER: PCT/US93/10555  
FILING DATE: 29-OCT-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Howells, Stacy L.  
REGISTRATION NUMBER: 34,842  
REFERENCE/DOCKET NUMBER: FD-2630  
TELEPHONE: (619) 455-5100  
TELEFAX: (619) 455-5110  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 120 amino acids

; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-428-197-22

Query Match 81.5%; Score 525; DB 5; Length 120;  
Best Local Similarity 84.2%; Pred. No. 1.5e-44;  
Matches 101; Conservative 6; Mismatches 13; Indels 0; Gaps 0;

QY 1 QVQLVQSGGGLVQPGGSLRLSCAASGFTFRSNAMGVWVRQAPGKGLWVSGISGSGSTYY 60  
||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
DB 1 QVQLVQSGGGLVQPGGSLRLSCAASGFTFRSNAMGVWVRQAPGKGLWVSDIASGSGSTYY 60  
||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
QY 61 ADSVKGRTTISRDNKNTLYLQMNLSRAEDTAVYYCAKHTGGGVMDPIDYWGQGTFLVTVS 120  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
DB 61 ADSVKGRTTISRDNKNTLYLQMNLSRAEDTALYYCASSGAGWGLPSLDYWGQGTFLVTVS 120  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

## RESULT 11

PCT-US93-10555-22

; Sequence 22, Application PC/TUS9310555

; GENERAL INFORMATION:

; APPLICANT: SILVERMAN, GREGG J.

; TITLE OF INVENTION: METHOD FOR STIMULATING PRODUCTION OF

; TITLE OF INVENTION: VARIABLE REGION GENE FAMILY RESTRICTED ANTIBODIES THROUGH

; TITLE OF INVENTION: VACCINATION WITH A B-CELL SUPERANTIGEN AND CONJUGATES

; TITLE OF INVENTION: THEREOF

; NUMBER OF SEQUENCES: 51

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Spensley Horn Jubas &amp; Lubitz

; STREET: 1880 Century Park East - Suite 500

; CITY: Los Angeles

; STATE: California

; COUNTRY: USA

; ZIP: 90067

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US93/10555

; FILING DATE: 29-OCT-1993

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: Howells, Stacy L.

; REGISTRATION NUMBER: 34,842

; REFERENCE/DOCKET NUMBER: FD-2630

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (619) 455-5100

; TELEFAX: (619) 455-5110

; INFORMATION FOR SEQ ID NO: 22:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 120 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

PCT-US93-10555-22

Query Match 81.5%; Score 525; DB 5; Length 120;  
Best Local Similarity 84.2%; Pred. No. 1.5e-44;  
Matches 101; Conservative 6; Mismatches 13; Indels 0; Gaps 0;

QY 1 QVQLVQSGGGLVQPGGSLRLSCAASGFTFRSNAMGVWVRQAPGKGLWVSGISGSGSTYY 60  
||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
DB 1 QVQLVQSGGGLVQPGGSLRLSCAASGFTFRSNAMGVWVRQAPGKGLWVSDIASGSGSTYY 60  
||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
QY 61 ADSVKGRTTISRDNKNTLYLQMNLSRAEDTAVYYCAKHTGGGVMDPIDYWGQGTFLVTVS 120  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
DB 61 ADSVKGRTTISRDNKNTLYLQMNLSRAEDTALYYCASSGAGWGLPSLDYWGQGTFLVTVS 120  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

## RESULT 12

US-08-983-607-36

; Sequence 36, Application US/08983607

; Patent No. 6140470

; GENERAL INFORMATION:

; APPLICANT: Xiahong Cai

; TITLE OF INVENTION: Human Anti-Tumor Monoclonal Anti-

; TITLE OF INVENTION: bodies

; NUMBER OF SEQUENCES: 51

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Department of Molecular Biophysics

; ADDRESSEE: and Biochemistry, Yale University

; STREET: 266 Whitney Avenue

; CITY: New Haven

; STATE: Connecticut

; COUNTRY: United States of America

; ZIP: 06520-8114

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5" 1.44 Mb diskette

; COMPUTER: IBM PC

; OPERATING SYSTEM: MS DOS

; SOFTWARE: Word Processing

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/983,607

; FILING DATE: April 27, 1998

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/IB96/01032

; FILING DATE: June 28, 1996

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Mary M. Krinsky

; REGISTRATION NUMBER: 32423

; REFERENCE/DOCKET NUMBER: OCR-679

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 203-773-9544

; TELEFAX: 203-773-1183

; INFORMATION FOR SEQ ID NO: 36:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 116 residues

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE:

; DESCRIPTION: polypeptide

; ORIGINAL SOURCE:

; ORGANISM: Homo sapiens (melanoma patient immu-

; ORGANISM: nized with autologous tumor cells)

; INDIVIDUAL ISOLATE: peripheral blood lymphocytes

; IMMEDIATE SOURCE:

; LIBRARY: DM414 scFv antibodies obtained from

; LIBRARY: fUSE5 fusion phage construct

; CLONE: V474

; FEATURE:

; NAME/KEY: heavy chain

; US-08-983-607-36

Query Match 81.4%; Score 524.5; DB 4; Length 116;  
Best Local Similarity 86.8%; Pred. No. 1.6e-44;  
Matches 105; Conservative 2; Mismatches 9; Indels 5; Gaps 2;

QY 1 QVQLVQSGGGLVQPGGSLRLSCAASGFTFRSNAMGVWVRQAPGKGLWVSGISGSGSTYY 60  
||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
DB 1 EVQLVQSGGGLVQPGGSLRLSCAASGFTFRSNAMGVWVRQAPGKGLWVSGISGSGSTYY 60  
||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
QY 61 ADSVKGRTTISRDNKNTLYLQMNLSRAEDTAVYYCAKHTGGGVMDPIDYWGQGTFLVTVS 120  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
DB 61 ADSVKGRTTISRDNKNTLYLQMNLSRAEDTAVYYCAKHTGGGVMDPIDYWGQGTFLVTVS 115  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
QY 121 S 121

Db 116 S 116

# RESULT 13

US-07-934-373C-21  
; Sequence 21, Application US/07934373C  
; Patent No. 5821337

## GENERAL INFORMATION:

APPLICANT: Paul J. Carter  
; APPLICANT: Leonard G. Presta  
; TITLE OF INVENTION: Immunoglobulin Variants  
; NUMBER OF SEQUENCES: 48  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 1 DNA Way  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080

## COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WinPatIn (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/934,373C  
; FILING DATE: 21-Aug-1992  
; CLASSIFICATION: 530

## PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US92/05126  
; FILING DATE: 15-JUN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/715272  
; FILING DATE: 14-JUN-1991

## ATTORNEY/AGENT INFORMATION:

NAME: Lee, Wendy M.  
; REGISTRATION NUMBER: 40,378  
; REFERENCE/DOCKET NUMBER: P0709P2

## TELECOMMUNICATION INFORMATION:

TELEPHONE: 650/225-1994  
; TELEFAX: 650/952-9881  
; INFORMATION FOR SEQ ID NO: 21:

## SEQUENCE CHARACTERISTICS:

LENGTH: 122 amino acids  
; TYPE: Amino Acid  
; TOPOLOGY: Linear

US-07-934-373C-21

Query Match 81.4%; Score 524.5; DB 2; Length 122;

Best Local Similarity 85.2%; Pred. No. 1.7e-44;

Matches 104; Conservative 4; Mismatches 13; Indels 1; Gaps 1;

Qy 1 QVQLVQSGGGLVQPGGSLRLSCAASGFTFRSNAMGWVRQAPGKLEWVSGISGSGSTYY 60

Db 1 EVQLVESGGGLVQPGGSLRLSCAASGFTFSYAMSWVRQAPGKLEWVSVISGDGSIYY 60

Qy 61 ADSVKGRFTISRDNKNTLYLQMSLRADTAIVYCAK-HTGGGVNDPDIYWGQGLTVTV 119

Db 61 ADSVKGRFTISRDNKNTLYLQMSLRADTAIVYCAK-HTGGGVNDPDIYWGQGLTVTV 120

Qy 120 SS 121

Db 121 SS 122

# RESULT 14

US-08-437-642B-21  
; Sequence 21, Application US/08437642B  
; Patent No. 6054297

## GENERAL INFORMATION:

APPLICANT: Paul J. Carter

APPLICANT: Leonard G. Presta  
; TITLE OF INVENTION: Immunoglobulin Variants  
; NUMBER OF SEQUENCES: 47  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 1 DNA Way  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080

## COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WinPatIn (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/437,642B  
; FILING DATE: 09-May-1995  
; CLASSIFICATION: 530

## PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/934373  
; FILING DATE: 21-AUG-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/146206  
; FILING DATE: 17-NOV-1993

## PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US92/05126  
; FILING DATE: 15-JUN-1992  
; APPLICATION NUMBER: 07/715272  
; FILING DATE: 14-JUN-1991

## ATTORNEY/AGENT INFORMATION:

NAME: Lee, Wendy M.  
; REGISTRATION NUMBER: 40,378  
; REFERENCE/DOCKET NUMBER: P0709P2C1

## TELECOMMUNICATION INFORMATION:

TELEPHONE: 650/225-1994  
; TELEFAX: 650/952-9881  
; INFORMATION FOR SEQ ID NO: 21:

## SEQUENCE CHARACTERISTICS:

LENGTH: 122 amino acids  
; TYPE: Amino Acid  
; TOPOLOGY: Linear

US-08-437-642B-21

Query Match 81.4%; Score 524.5; DB 3; Length 122;

Best Local Similarity 85.2%; Pred. No. 1.7e-44;

Matches 104; Conservative 4; Mismatches 13; Indels 1; Gaps 1;

Qy 1 QVQLVQSGGGLVQPGGSLRLSCAASGFTFRSNAMGWVRQAPGKLEWVSGISGSGSTYY 60

Db 1 EVQLVESGGGLVQPGGSLRLSCAASGFTFSYAMSWVRQAPGKLEWVSVISGDGSIYY 60

Qy 61 ADSVKGRFTISRDNKNTLYLQMSLRADTAIVYCAK-HTGGGVNDPDIYWGQGLTVTV 119

Db 61 ADSVKGRFTISRDNKNTLYLQMSLRADTAIVYCAK-HTGGGVNDPDIYWGQGLTVTV 120

Qy 120 SS 121

Db 121 SS 122

# RESULT 15

PCT-US93-07832-21

; Sequence 21, Application PC/TUS9307832

## GENERAL INFORMATION:

APPLICANT: Genentech, Inc.  
; TITLE OF INVENTION: Immunoglobulin Variants  
; NUMBER OF SEQUENCES: 40  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 460 Point San Bruno Blvd

```

; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/07832
; FILING DATE: 19930820
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/715272
; FILING DATE: 14-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/05126
; FILING DATE: 15-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/934373
; FILING DATE: 21-AUG-1992
; ATTORNEY/AGENT INFORMATION:
; NAME:
; REGISTRATION NUMBER:
; REFERENCE/DOCKET NUMBER: 709P2PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE:
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 122 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; PCT-US93-07832-21

```

```

Query Match      81.4%; Score 524.5; DB 5; Length 122;
Best Local Similarity 85.2%; Pred. No. 1.7e-44;
Matches 104; Conservative 4; Mismatches 13; Indels 1; Gaps 1;

QY 1 QVQLVQSGGGLVQPGGSLRLSCAASGFTFRSNAMGWVQAQPKGLEWVSIGSGGSTYY 60
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 EVQLVESGGGLVQPGGSLRLSCAASGFTFSSYAMSVWQAQPKGLEWVSISDGGSTYY 60
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 61 ADSVKGRTISRDNKNTLYLQMSLRAEDTAVYYCAK-HTGGGVWDPIDYWGQGLTVTV 119
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 ADSVKGRTISRDNKNTLYLQMSLRAEDTAVYYCARGVYSLSGLDYWGQGLTVTV 120
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 120 SS 121
   ||
Db 121 SS 122

```

Search completed: October 4, 2002, 10:22:14  
Job time: 162 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 4, 2002, 10:20:33 ; Search time 24.92 Seconds  
(without alignments)  
466.565 Million cell updates/sec

Title: US-09-822-698A-3  
perfect score: 644  
Sequence: 1 QVQLVSGGGLVQPGGSLRL.....GGVWDPTDYWGQGTLVTVSS 121

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 s

```
Database :      PIR_71:*
1:  pir1:*
2:  pir2:*
3:  pir3:*
4:  pir4:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query %		Length	DB	ID	Description
		Match					
1	543	84.3	119	2	S31107	Ig heavy chain - h	
2	542	84.2	119	2	S36005	Ig heavy chain v r	
3	535.5	83.2	120	2	S48798	Ig heavy chain v r	
4	534	82.9	140	2	S31686	Ig heavy chain v r	
5	531	82.5	119	2	S31108	Ig heavy chain - h	
6	528	82.0	119	2	S36005	Ig heavy chain - h	
7	528	82.0	121	2	D55673	Ig heavy chain v r	
8	527	81.8	127	2	S38489	Ig heavy chain - h	
9	527	81.8	140	2	S31588	Ig heavy chain v r	
10	526	81.7	134	2	S31699	Ig heavy chain v r	
11	522	81.1	123	2	S31114	Ig heavy chain - h	
12	522	81.1	138	2	S31666	Ig heavy chain v r	
13	519	80.6	160	2	S05271	Ig heavy chain pre	
14	518.5	80.5	124	2	S20782	Ig heavy chain v r	
15	518	80.4	120	2	S38278	Ig heavy chain v r	
16	509.5	79.1	112	2	PH1647	Ig heavy chain v r	
17	508.5	79.0	116	2	S31110	Ig heavy chain - h	
18	503	78.1	109	2	PH1649	Ig heavy chain v r	
19	501	77.8	121	2	S19666	Ig heavy chain v r	
20	501	77.8	140	2	S40532	Ig heavy chain v r	
21	498	77.3	121	2	G36005	Ig heavy chain pre	
22	495.5	76.9	108	2	PH1648	Ig heavy chain v r	
23	494	76.7	134	2	S31679	Ig heavy chain v r	
24	492.5	76.5	122	2	S20772	Ig heavy chain v r	
25	492	76.4	121	2	S31113	Ig heavy chain - h	
26	492	76.4	143	2	S23624	Ig heavy chain v r	
27	490	76.1	135	2	S31598	Ig heavy chain v r	
28	489.5	76.0	120	2	S44111	Ig heavy chain v-D	
29	489.5	76.0	151	2	A60943	Ig heavy chain pre	

A:Cross-references: GB:M18513

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F:15-98/Domain: immunoglobulin homology <IMM>

Query Match 84.2%; Score 542; DB 2; Length 119;  
Best Local Similarity 87.6%; Pred. No. 2.4e-40;  
Matches 106; Conservative 3; Mismatches 10; Indels 2; Gaps 1;  
QY 1 QVQLVQSGGGLVQPGGSLRLSCAASGFTFRSNAMGWVRQAPGKLEWVSGISGGSTYY 60  
Db 1 EVQLLESGGGLVQPGGSLRLSCAASGFTFRSNAMGWVRQAPGKLEWVSAISGGSTYY 60  
QY 61 ADSVKGRFTISRDNKNTLYQMNSLRADTAIVYCAKHTGGVMDPIDYWGQGLTVTVS 120  
Db 61 ADSVKGRFTISRDNKNTLYQMNSLRADTAIVYCAKHTGGVMDPIDYWGQGLTVTVS 118  
QY 121 S 121  
Db 119 S 119

#### RESULT 3

S48798

Ig heavy chain V region (anti-Sm, VH3/Dxp4/JH4b) - human

C:Species: Homo sapiens (man)

C>Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 23-Jul-1999

C:Accession: S48798

R:Mamoudi, M.; Edwards, J.; Cairns, E.; Bell, D.

submitted to the EMBL Data Library, October 1994

A:Description: Molecular characterization of natural human anti-Sm autoantibodies.

A:Reference number: S48797

A:Accession: S48798

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-120 <NA>

A:Cross-references: EMBL:Z46382; NID:G562324; PIDN:CAR86521.1; PID:G1340167

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F:15-98/Domain: immunoglobulin homology <IMM>

Query Match 83.2%; Score 535.5; DB 2; Length 120;  
Best Local Similarity 86.9%; Pred. No. 8.9e-40;  
Matches 106; Conservative 3; Mismatches 10; Indels 3; Gaps 2;  
QY 1 QVQLVQSGGGLVQPGGSLRLSCAASGFTFRSNAMGWVRQAPGKLEWVSGISGGSTYY 60  
Db 1 EVQLLESGGGLVQPGGSLRLSCAASGFTFRSNAMGWVRQAPGKLEWVSAISGGSTYY 60  
QY 61 ADSVKGRFTISRDNKNTLYQMNSLRADTAIVYCAKHTGGVMDPI-DYWGQGLTVTV 119  
Db 61 ADSVKGRFTISRDNKNTLYQMNSLRADTAIVYCAKHTGGVMDPI-DYWGQGLTVTV 118  
QY 120 SS 121  
Db 119 SS 120

#### RESULT 4

S31886

Ig heavy chain V region - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 22-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 23-Jul-1999

C:Accession: S31886

R:Cuisinier, A.M.; Gauthier, L.; Boubli, L.; Fougereau, M.; Tonnelle, C.

submitted to the EMBL Data Library, June 1992

A:Description: Mechanisms that generate human immunoglobulin diversity operate from the

A:Reference number: S31585

A:Accession: S31886

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-140 <CUI>

A:Cross-references: EMBL:Z14205; NID:G30969; PIDN:CAA74574.1; PID:G30970

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F:34-117/Domain: immunoglobulin homology <IMM>

Query Match 82.9%; Score 534; DB 2; Length 140;  
Best Local Similarity 86.0%; Pred. No. 1.4e-39;  
Matches 104; Conservative 4; Mismatches 13; Indels 0; Gaps 0;  
QY 1 QVQLVQSGGGLVQPGGSLRLSCAASGFTFRSNAMGWVRQAPGKLEWVSGISGGSTYY 60  
Db 20 EVQLLESGGGLVQPGGSLRLSCAASGFTFRSNAMGWVRQAPGKLEWVSAISGGSTYY 79  
QY 61 ADSVKGRFTISRDNKNTLYQMNSLRADTAIVYCAKHTGGVMDPIDYWGQGLTVTVS 120  
Db 80 SDSVKGRFTISRDNKNTLYQMNSLRADTAIVYCAKHTGGVMDPIDYWGQGLTVTVS 139  
QY 121 S 121  
Db 140 S 140

#### RESULT 5

S31108

Ig heavy chain - human

C:Species: Homo sapiens (man)

C>Date: 02-Dec-1993 #sequence\_revision 26-May-1995 #text\_change 17-Mar-1999

C:Accession: S31108

R:Raaphorst, F.M.; Timmers, E.; Kenter, M.J.H.; van Tol, M.J.D.; Vossen, J.M.; Schuur

Eur. J. Immunol. 22, 247-251, 1992

A:Title: Restricted utilization of germ-line V(H)3 genes and short diverse third comp

A:Reference number: S31104; MUID:92111633

A:Accession: S31108

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: mRNA

A:Residues: 1-119 <RAA>

A:Cross-references: EMBL:X62956

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1991

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F:15-98/Domain: immunoglobulin homology <IMM>

Query Match 82.5%; Score 531; DB 2; Length 119;  
Best Local Similarity 85.5%; Pred. No. 2.2e-39;  
Matches 106; Conservative 3; Mismatches 7; Indels 8; Gaps 2;  
QY 1 QVQLVQSGGGLVQPGGSLRLSCAASGFTFRSNAMGWVRQAPGKLEWVSGISGGSTYY 60  
Db 1 EVQLLESGGGLVQPGGSLRLSCAASGFTFRSNAMGWVRQAPGKLEWVSAISGGSTYY 60  
QY 61 ADSVKGRFTISRDNKNTLYQMNSLRADTAIVYCAKHTGGVMDPIDYWGQGLTV 117  
Db 61 ADSVKGRFTISRDNKNTLYQMNSLRADTAIVYCAKHTGGVMDPIDYWGQGLTV 115  
QY 118 TVSS 121  
Db 116 TVSS 119

#### RESULT 6

D36005

Ig heavy chain V region (M43) - human

C:Species: Homo sapiens (man)

C>Date: 21-Dec-1990 #sequence\_revision 21-Dec-1990 #text\_change 16-Dec-1998

C:Accession: D36005

R:Schroeder Jr., H.W.; Wang, J.Y.

Proc. Natl. Acad. Sci. U.S.A. 87, 6146-6150, 1990

A:Title: Preferential utilization of conserved immunoglobulin heavy chain variable ge

A:Reference number: A36005; MUID:90349571

A:Accession: D36005

A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-119 <SCH>  
 A:Cross-references: GB:M34024  
 C:Genetics:  
 A:Gene: GDB:IGH4; IGHDI1  
 A:Cross-references: GDB:1118731; OMIM:146910  
 A:Map position: 14q32.33-14q32.33  
 C:Superfamily: immunoglobulin V region; immunoglobulin homology  
 C:Keywords: heterotetramer; immunoglobulin  
 F:15-98/Domain: immunoglobulin homology <IMM>

Query Match 82.0%; Score 528; DB 2; Length 119;  
 Best Local Similarity 86.0%; Pred. No. 4e-39;  
 Matches 104; Conservative 3; Mismatches 12; Indels 2; Gaps 1;

QY 1 QVQLVQSGGGLVQPGGSLRLSCAASGFTFRSNAMGWVRQAPGKGLWVSGISGGSTYY 60  
 :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 Db 1 EVQLLESGGGLVQPGGSLRLSCAASGFTFSYAMSVWRQAPGKGLWVSAISGGSTYY 60  
 :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 QY 61 ADSVKGRFTISRDNKNTLYQMNSLAEDTAVYYCAKHTGGGVWDPIYWGQGLTVTVS 120  
 :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 Db 61 ADSVKGRFTISRDNKNTLYQMNSLAEDTAVYYCAKDKDNDW--FDPWGQGLTVTVS 118  
 :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 QY 121 S 121  
 Db 119 S 119

## RESULT 7

I55673  
 Ig heavy chain - human (fragment)  
 C:Species: Homo sapiens (man)  
 C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 21-Jan-2000  
 C:Accession: I55673  
 R:Knight, G.B.; Agnello, V.; Bonagura, V.; Barnes, J.L.; Panka, D.J.; Zhang, Q.X.  
 J. Exp. Med. 178, 1903-1911, 1993  
 A:Title: Human rheumatoid factor cross-idiotypes. IV. Studies on WA Xid-positive IgM with tinct from the 17.109 and G6 Xids.  
 A:Reference number: I55673; MUID:9406558  
 A:Accession: I55673  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-121 <RES>  
 A:Cross-references: GB:M87268; NID:g186197; PIDN:AAC37536.1; PID:g186198  
 C:Genetics:  
 A:Gene: GDB:IGHM  
 A:Cross-references: GDB:120086; OMIM:147020  
 A:Map position: 14q32.33-14q32.33  
 C:Superfamily: immunoglobulin V region; immunoglobulin homology  
 F:15-98/Domain: immunoglobulin homology <IMM>

Query Match 82.0%; Score 528; DB 2; Length 121;  
 Best Local Similarity 83.1%; Pred. No. 4e-39;  
 Matches 103; Conservative 5; Mismatches 10; Indels 6; Gaps 2;

QY 1 QVQLVQSGGGLVQPGGSLRLSCAASGFTFRSNAMGWVRQAPGKGLWVSGISGGSTYY 60  
 :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 Db 1 EVQLLESGGGLVQPGGSLRLSCTASGFTSTYGNMSVWRQAPGKGLWVSAISGGSTYY 60  
 :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 QY 61 ADSVKGRFTISRDNKNTLYQMNSLAEDTAVYYCA--KHTGGGVWDPIYWGQGLTV 117  
 :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 Db 61 ADSVKGRFTISRDNKNTLYQMNSLAEDTAVYYCAAPRHAGS---PPYDIWGQGLTV 117  
 :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 QY 118 TVSS 121  
 Db 118 TVSS 121

## RESULT 8

S38489

Ig heavy chain - human (fragment)  
 C:Species: Homo sapiens (man)  
 C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 23-Jul-1999  
 C:Accession: S38489  
 R:Marks, J.D.; Ouwehand, W.H.; Bye, J.M.; Pinner, R.; Jorick, B.D.; Voak, D.; Thorpe  
 submitted to the EMBL Data Library, June 1993  
 A:Description: Human antibody fragments specific for human blood group antigens from  
 A:Reference number: S38488  
 A:Accession: S38489  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-127 <MAR>  
 A:Cross-references: EMBL:Z23028; NID:g414025; PIDN:CAAG0563.1; PID:g414026  
 C:Superfamily: immunoglobulin V region; immunoglobulin homology  
 C:Keywords: heterotetramer; immunoglobulin  
 F:15-98/Domain: immunoglobulin homology <IMM>

Query Match 81.8%; Score 527; DB 2; Length 127;  
 Best Local Similarity 82.0%; Pred. No. 5.2e-39;  
 Matches 105; Conservative 3; Mismatches 10; Indels 10; Gaps 2;

QY 1 QVQLVQSGGGLVQPGGSLRLSCAASGFTFRSNAMGWVRQAPGKGLWVSGISGGSTYY 60  
 :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 Db 1 QVQLVQSGGGLVQPGGSLRLSCAASGFTFSYAMSVWRQAPGKGLWVSAISGGSTYY 60  
 :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 QY 61 ADSVKGRFTISRDNKNTLYQMNSLAEDTAVYYCAK-----HTGGGVWDPIYW 111  
 :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 Db 61 ADSVKGRFTISRDNKNTLYQMNSLAEDTAVYYCAKEGPTFA-DYDSSGYIS-FDYW 119  
 :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 QY 112 QGGLTVTV 119  
 Db 120 QGGLTVTV 127

## RESULT 9

S31588  
 Ig heavy chain V region - human (fragment)  
 C:Species: Homo sapiens (man)  
 C:Date: 22-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 23-Jul-1999  
 C:Accession: S31588  
 R:Cuisinier, A.M.; Gauthier, L.; Boubli, L.; Fougereau, M.; Tonnel, C.  
 submitted to the EMBL Data Library, June 1992  
 A:Description: Mechanisms that generate human immunoglobulin diversity operate from t  
 A:Reference number: S31585  
 A:Accession: S31588  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-140 <CUI>  
 A:Cross-references: EMBL:Z14200; NID:g30957; PIDN:CAA4369.1; PID:g30958  
 C:Superfamily: immunoglobulin V region; immunoglobulin homology  
 C:Keywords: heterotetramer; immunoglobulin  
 F:34-117/Domain: immunoglobulin homology <IMM>

Query Match 81.8%; Score 527; DB 2; Length 140;  
 Best Local Similarity 85.1%; Pred. No. 5.7e-39;  
 Matches 103; Conservative 5; Mismatches 13; Indels 0; Gaps 0;

QY 1 QVQLVQSGGGLVQPGGSLRLSCAASGFTFRSNAMGWVRQAPGKGLWVSGISGGSTYY 60  
 :||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 Db 20 EVQLLESGGGLVQPGGSLRLSCAASGFTFSYAMSVWRQAPGKGLWVSAISGGSTYY 79  
 :||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 QY 61 ADSVKGRFTISRDNKNTLYQMNSLAEDTAVYYCAKHTGGGVWDPIYWGQGLTVTVS 120  
 :||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 Db 80 ADSVKGRFTISRDNKNTLYQMNSLAEDTAVYYCAKDHDI-SNLYFDYWGQGLTVTVS 139  
 :||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 QY 121 S 121  
 Db 140 S 140

## RESULT 10

S31699

Ig heavy chain V region - human (fragment)  
 C:Species: Homo sapiens (man)  
 C:Date: 22-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 23-Jul-1999  
 C:Accession: S31699  
 R:Cuissinier, A.M.; Gauthier, L.; Boubli, L.; Fougereau, M.; Tonnelie, C.  
 submitted to the EMBL Data Library, June 1992  
 A:Description: Mechanisms that generate human immunoglobulin diversity operate from the  
 A:Reference number: S31585  
 A:Accession: S31699  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-134 <CUI>  
 A:Cross-references: EMBL:214201; NID:g30961; PIDN:CAA78570.1; PID:g30962  
 C:Superfamily: immunoglobulin V region; immunoglobulin homology  
 C:Keywords: heterotetramer; immunoglobulin  
 F:34-117/Domain: immunoglobulin homology <IMM>

Query Match 81.7%; Score 526; DB 2; Length 134;  
 Best Local Similarity 83.5%; Pred. No. 6.7e-39;  
 Matches 101; Conservative 6; Mismatches 8; Indels 6; Gaps 1;  
 QY 1 QVOLVSGGGLVOPGGSLRLSCAASGFTFRSNAMGWVRQAPGKLEWVSGISGGSTYY 60  
 :||||:||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 Db 20 EVQLLESGGGLVOPGGSLRLSCAASGFTFRSNAMGWVRQAPGKLEWVSAISGGSTYY 79  
 :||||:||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 QY 61 ADSVKGRFTISRDNKNTLYLQMSLRADTAIVYCAKHTGGVWDPIDYWGOGTLTVTS 120  
 :||||:||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 Db 80 SDVKGRFTISRDNKNTLYLQMSLRADTAIVYCAR-----WRDLTDYWGOGTLTVTS 133  
 QY 121 S 121  
 Db 134 S 134

RESULT 11

S31114  
 Ig heavy chain - human  
 C:Species: Homo sapiens (man)  
 C:Date: 02-Dec-1993 #sequence\_revision 26-May-1995 #text\_change 17-Mar-1999  
 C:Accession: S31114  
 R:Raaphorst, F.M.; Timmers, E.; Kenter, M.J.H.; van Tol, M.J.D.; Vossen, J.M.; Schuurman  
 Eur. J. Immunol. 22, 247-251, 1992  
 A:Title: Restricted utilization of germ-line V(H)3 genes and short diverse third complement  
 A:Reference number: S31104; MUID:92111633  
 A:Accession: S31114  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-123 <RAA>  
 A:Cross-references: EMBL:X62963  
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1991  
 C:Superfamily: immunoglobulin V region; immunoglobulin homology  
 C:Keywords: heterotetramer; immunoglobulin  
 F:15-98/Domain: immunoglobulin homology <IMM>

Query Match 81.1%; Score 522; DB 2; Length 123;  
 Best Local Similarity 84.7%; Pred. No. 1.4e-38;  
 Matches 105; Conservative 4; Mismatches 11; Indels 4; Gaps 2;  
 QY 1 QVOLVSGGGLVOPGGSLRLSCAASGFTFRSNAMGWVRQAPGKLEWVSGISGGSTYY 60  
 :||||:||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 Db 1 EVQLLESGGGLVOPGGSLRLSCAASGFTFRSNAMGWVRQAPGKLEWVSAISGGSTYY 60  
 QY 61 ADSVKGRFTISRDNKNTLYLQMSLRADTAIVYCAKHTGGVWDPIDYWGOGTLV 117  
 :||||:||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 Db 61 ADSVKGRFTISRDNKNTLYLQMSLRADTAIVYCARSLYLFLEW-LFDYWGOGTLV 119  
 QY 118 TVSS 121  
 Db 120 TVSS 123

RESULT 12

S31666  
 Ig heavy chain V region - human (fragment)  
 C:Species: Homo sapiens (man)  
 C:Date: 22-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 23-Jul-1999  
 C:Accession: S31666  
 R:Cuissinier, A.M.; Gauthier, L.; Boubli, L.; Fougereau, M.; Tonnelie, C.  
 submitted to the EMBL Data Library, June 1992  
 A:Description: Mechanisms that generate human immunoglobulin diversity operate from t  
 A:Reference number: S31585  
 A:Accession: S31666  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-138 <CUI>  
 A:Cross-references: EMBL:214202; NID:g30963; PIDN:CAA78571.1; PID:g30964  
 C:Superfamily: immunoglobulin V region; immunoglobulin homology  
 C:Keywords: heterotetramer; immunoglobulin  
 F:34-117/Domain: immunoglobulin homology <IMM>

Query Match 81.1%; Score 522; DB 2; Length 138;  
 Best Local Similarity 86.0%; Pred. No. 1.5e-38;  
 Matches 104; Conservative 4; Mismatches 11; Indels 2; Gaps 2;  
 QY 1 QVOLVSGGGLVOPGGSLRLSCAASGFTFRSNAMGWVRQAPGKLEWVSGISGGSTYY 60  
 :||||:||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 Db 20 EVQLLESGGGLVOPGGSLRLSCAASGFTFRSNAMGWVRQAPGKLEWVSAISGGSTYY 79  
 :||||:||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 QY 61 ADSVKGRFTISRDNKNTLYLQMSLRADTAIVYCAKHTGGVWDPIDYWGOGTLTVTS 120  
 :||||:||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 Db 80 ADSVKGRFTISRDNKNTLYLQMSLRADTAIVYCAKHTGGVWDPIDYWGOGTLTVTS 137  
 QY 121 S 121  
 Db 138 S 138

RESULT 13

S05271  
 Ig heavy chain precursor - human (fragment)  
 C:Species: Homo sapiens (man)  
 C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 16-Aug-1996  
 C:Accession: S05271; S04602  
 R:Kishimoto, T.  
 submitted to the EMBL Data Library, March 1989  
 A:Reference number: S05270  
 A:Accession: S05271  
 A:Molecule type: mRNA  
 A:Residues: 1-160 <KIS1>  
 A:Cross-references: EMBL:X14584  
 R:Kishimoto, T.; Okajima, H.; Okumoto, T.; Taniguchi, M.  
 Nucleic Acids Res. 17, 4385, 1989  
 A:Title: Nucleotide sequences of the cDNAs encoding L- and V-regions of H- and L-chains  
 A:Reference number: S04601; MUID:89296497  
 A:Accession: S04602  
 A:Molecule type: mRNA  
 A:Residues: 1-144 <KIS2>  
 A:Cross-references: EMBL:X14584  
 C:Superfamily: immunoglobulin V region; immunoglobulin homology  
 C:Keywords: heterotetramer; immunoglobulin  
 F:1-19/Domain: signal sequence #status predicted <SIG>  
 F:20-160/Product: Ig heavy chain (fragment) #status predicted <MAT>  
 F:34-117/Domain: immunoglobulin homology <IMM>

Query Match 80.6%; Score 519; DB 2; Length 160;  
 Best Local Similarity 82.4%; Pred. No. 3.3e-38;  
 Matches 103; Conservative 6; Mismatches 12; Indels 4; Gaps 1;  
 QY 1 QVOLVSGGGLVOPGGSLRLSCAASGFTFRSNAMGWVRQAPGKLEWVSGISGGSTYY 60  
 :||||:||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 Db 20 EVQLLESGGGLVOPGGSLRLSCAASGFTFRSNAMGWVRQAPGKLEWVSAISGGSTYY 79



Q/ 61 ADSVKGRTISRDNKNTLYLQMSLRAEDTAVYYCAKHTGGV----WDPIDYWGQGTLL 116  
|||||  
Db 80 ADSVKGRTISRDNKNTLYLQMSLRAEDTAVYYCAKAVRGVSYYYIGNDYWGQGT 139  
|||||  
QY 117 VTVSS 121  
|||||  
Db 140 VTVSS 144  
|||||

## RESULT 14

S20782  
Ig heavy chain V region - human  
C:Species: Homo sapiens (man)  
C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 23-Mar-2001  
C:Accession: S20782  
R:Wortari, F.; Wang, J.; Schroeder, H.W.  
Submitted to the EMBL Data Library, April 1992  
A:Description: Analysis of the IgA and IgG rearranged VH repertoire of human cord blood  
A:Reference number: S20765  
A:Accession: S20782  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-124 <OR>  
A:Cross-references: EMBL:Z11946; NID:g33897; PIDN:CRA78003.1; PID:g33898  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotetramer; immunoglobulin  
F:15-98/Domain: immunoglobulin homology <IM>

Query Match 80.5%; Score 518.5; DB 2; Length 124;  
Best Local Similarity 83.9%; Pred. No. 2.8e-38;  
Matches 104; Conservative 4; Mismatches 13; Indels 3; Gaps 1;

QY 1 QVQLVSGGGLVPGGSLRLSCAASGTFRRSNAMGWVRQAPGKLEWVSGISGSGSTYY 60  
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 1 EVQLLEGGGGLVPGGSLRLSCAASGTFSSYAMNVRQAPGKLEWVSTISGSGDSTYY 60  
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
QY 61 ADSVKGRTISRDNKNTLYLQMSLRAEDTAVYYCAKH---TGGGVNDPDIYWGQGTLY 117  
|||||  
Db 61 ADSVKGRTISRDNKNTLYLQMSLRAEDTAVYYCAKERIAIFGVVIPHEDYWGQGTLY 120  
|||||  
QY 118 TVSS 121  
|||||  
Db 121 TVSS 124  
|||||

## RESULT 15

S36278  
Ig heavy chain V region (clone alpha-THY-23) - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 03-Feb-1994 #sequence\_revision 03-Feb-1994 #text\_change 23-Jul-1999  
C:Accession: S36278  
R:Griffiths, A.D.; Malmqvist, M.; Marks, J.D.; Bye, J.M.; Embleton, M.J.; McCafferty, J.  
EMBO J. 12, 725-734, 1993  
A:Title: Human anti-self antibodies with high specificity from phage display libraries.  
A:Reference number: S36256; MUID:93178448  
A:Accession: S36278  
A:Status: preliminary; nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-120 <GRI>  
A:Cross-references: EMBL:Z18830; NID:g33114; PIDN:CRA79282.1; PID:g939894  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotetramer; immunoglobulin  
F:15-98/Domain: immunoglobulin homology <IM>

Query Match 80.4%; Score 518; DB 2; Length 120;  
Best Local Similarity 85.8%; Pred. No. 3e-38;  
Matches 103; Conservative 3; Mismatches 14; Indels 0; Gaps 0;

QY 1 QVQLVSGGGLVPGGSLRLSCAASGTFRRSNAMGWVRQAPGKLEWVSGISGSGSTYY 60  
|||||  
Db 1 EVQLLEGGGGLVPGGSLRLSCAASGTFSSYAMNVRQAPGKLEWVSTISGSGDSTYY 60  
|||||

Db 1 QVQLVSGGGLVPGGSLRLSCAASGTFRRSNAMGWVRQAPGKLEWVSGISGSGSTYY 60  
QY 61 ADSVKGRTISRDNKNTLYLQMSLRAEDTAVYYCAKHTGGVNDPDIYWGQGTLYTVTS 120  
|||||  
Db 61 ADSVKGRTISRDNKNTLYLQMSLRAEDTAVYYCAKGMTVAAYFYEDYWGQGTLYTVTS 120  
|||||

Search completed: October 4, 2002, 10:26:16  
Job time: 343 sec

10/4/2002 11:00:31 AM

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 4, 2002, 10:22:18 ; Search time 15.77 seconds  
(without alignments)  
297.087 Million cell updates/sec

Title: US-09-822-698A-3  
Perfect score: 644  
Sequence: 1 QVQLVQSGGGLVQPGGSLRL.....GGVMDPIDYWGQGLTVTVSS 121

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40.\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	471	73.1	117	1 HV3C_HUMAN	P01764 homo sapien
2	466.5	72.4	114	1 HV3B_HUMAN	P01763 homo sapien
3	456.5	70.9	122	1 HV3G_HUMAN	P01768 homo sapien
4	450.5	70.0	122	1 HV3A_HUMAN	P01768 homo sapien
5	449.5	69.8	116	1 HV3T_HUMAN	P01781 homo sapien
6	447	69.4	121	1 HV3J_HUMAN	P01771 homo sapien
7	438.5	68.1	116	1 HV05_CARAU	P19181 carassius a
8	437.5	67.9	122	1 HV3H_HUMAN	P01772 homo sapien
9	437.5	67.9	126	1 HV3K_HUMAN	P01772 homo sapien
10	435.5	67.6	136	1 HV16_MOUSE	P01783 mus musculus
11	435	67.5	115	1 HV3D_HUMAN	P01765 homo sapien
12	432	67.1	117	1 HV02_CANFA	P01785 canis fami
13	428	66.5	119	1 HV31_HUMAN	P01770 homo sapien
14	427	66.3	115	1 HV3F_HUMAN	P01767 homo sapien
15	422.5	65.6	120	1 HV3U_HUMAN	P01782 homo sapien
16	419.5	65.1	119	1 HV3M_HUMAN	P01774 homo sapien
17	419	65.1	120	1 HV3E_HUMAN	P01766 homo sapien
18	417.5	64.8	119	1 HV3N_HUMAN	P01775 homo sapien
19	414	64.3	119	1 HV3L_HUMAN	P01773 homo sapien
20	400.5	62.2	114	1 HV01_CANFA	P01784 canis fami
21	398	61.8	117	1 HV54_MOUSE	P18525 mus musculus
22	397.5	61.7	117	1 HV3Q_HUMAN	P01776 homo sapien
23	396.5	61.6	97	1 HV56_MOUSE	P18527 mus musculus
24	396	61.5	117	1 HV55_MOUSE	P18526 mus musculus
25	395.5	61.4	119	1 HV3P_HUMAN	P01777 homo sapien
26	395.5	61.4	122	1 HV20_MOUSE	P01789 mus musculus
27	395	61.3	115	1 HV32_MOUSE	P01801 mus musculus
28	394.5	61.3	122	1 HV21_MOUSE	P01790 mus musculus
29	394	61.2	116	1 HV3Q_HUMAN	P01778 homo sapien
30	393.5	61.1	111	1 HV35_MOUSE	P01804 mus musculus
31	393	61.0	113	1 HV30_MOUSE	P01799 mus musculus
32	390	60.6	142	1 HV01_RAT	P01805 rattus norv
33	389	60.4	116	1 HV3R_HUMAN	P01779 homo sapien

34	388	60.2	113	1 HV27_MOUSE	P01796 mus musculus
35	387.5	60.2	119	1 HV37_MOUSE	P01807 mus musculus
36	386.5	60.0	119	1 HV40_MOUSE	P01810 mus musculus
37	384.5	59.7	119	1 HV38_MOUSE	P01808 mus musculus
38	384	59.6	113	1 HV31_MOUSE	P01800 mus musculus
39	384	59.6	115	1 HV33_MOUSE	P01802 mus musculus
40	383	59.5	123	1 HV18_MOUSE	P01787 mus musculus
41	382	59.3	113	1 HV28_MOUSE	P01797 mus musculus
42	382	59.3	123	1 HV19_MOUSE	P01788 mus musculus
43	380	59.0	98	1 HV57_MOUSE	P18528 mus musculus
44	380	59.0	123	1 HV25_MOUSE	P01794 mus musculus
45	379	58.9	123	1 HV23_MOUSE	P01792 mus musculus

## ALIGNMENTS

```

RESULT 1
HV3C_HUMAN
ID HV3C_HUMAN STANDARD; PRT; 117 AA.
AC P01764;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE Iq heavy chain V-III region VH26 precursor.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=81101090; PubMed=6450418;
RA Matthysens G., Rabbitts T.H.;
RT "Structure and multiplicity of genes for the human immunoglobulin
heavy chain variable region.";
RL Proc. Natl. Acad. Sci. U.S.A. 77:6561-6565(1980).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; J00236; AAA53516.1; -
CC EMBL; M35415; AAA58735.1; -
CC PIR; A02047; H3HU26.
CC InterPro; IPR003006; Iq_MHC.
CC InterPro; IPR003596; Iq_v.
CC Pfam; PF00047; Iq; 1.
CC SMART; SM00406; Iq; 1.
CC Immunoglobulin V region; Signal.
FT SIGNAL
FT CHAIN 1 19
FT NON_TER 117 117 IG HEAVY CHAIN V-III REGION VH26.
SQ SEQUENCE 117 AA; 12582 MW; E826733F1A3CB0F1 CRC64;

```

```

Query Match 73.1%; Score 471; DB 1; Length 117;
Best Local Similarity 91.8%; Pred. No. 7.7e-40;
Matches 90; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
QY 1 QVQLVQSGGGLVQPGGSLRLSCAASGFTFRSNAMGWVRQAPGKGLVWVSGISGSGSTYY 60
:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 20 EVQLLESGGGLVQPGGSLRLSCAASGFTFSYSSNWVRQAPGKGLVWVSAISGSGSTYY 79
:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 61 ADSVKGRTISRDNSKNTLYLQMNLSRAEDTAVYYCAK 98
:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 80 GDSVKGRTISRDNSKNTLYLQMNLSRAEDTAVYYCAK 117
:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

```

RESULT 2

```
HV3B_HUMAN
ID HV3B_HUMAN STANDARD; PRT; 114 AA.
AC P01763;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE Ig heavy chain V-III region WEA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=83273707; PubMed=6410398;
RA Goni F., Frangione B.;
RT "Amino acid sequence of the Fv region of a human monoclonal IgM
RT (protein WEA) with antibody activity against 3,4-pyruvylated
RT galactose in Klebsiella polysaccharides K30 and K33.";
RL Proc. Natl. Acad. Sci. U.S.A. 80:4837-4841(1983).
CC -1- MISCELLANEOUS: THIS CHAIN WAS OBTAINED FROM A MONOCLONAL ANTIBODY
CC AGAINST 3,4-PYRUVYLATED GALACTOSE AND ISOLATED FROM A PATIENT WITH
CC WALDENSTROM'S MACROGLOBULINEMIA.
DR PTR; A02046; M3HUME.
DR HSSP; P01772; 2IG2.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig_1.
DR SMART; SM00406; Igv; 1.
KW Immunoglobulin V region.
FT MOD_RES 1
FT NON_TER 114 114 PYRROLIDONE CARBOXYLIC ACID.
SQ SEQUENCE 114 AA; 12256 MW; D88294FB418A07B7 CRC64;

Query Match 72.4%; Score 466.5; DB 1; Length 114;
Best Local Similarity 76.0%; Pred. No. 2.1e-39;
Matches 92; Conservative 11; Mismatches 11; Indels 7; Gaps 2;

QY 1 QVQLVSGGGLVQPGGSLRLSCAASGFTFRSNAMGWVRQAPGKGLVWVSIGSGGSTYY 60
DB 1 QVQLVSGGGLVQPGGSLRLSCAASGFTFRSNAMGWVRQAPGKGLVWVSIGSGGSTYY 60
QY 61 ADSVKGRFTISRDNKNTLYLQMSLRADTAIVYCAKHTGGVWDPIDYWGQGLTVTVS 120
DB 61 ADSVKGRFTISRDNKNTLYLQMSLRADTAIVYCAKHTGGVWDPIDYWGQGLTVTVS 120
QY 121 S 121
DB 114 S 114

RESULT 3
HV3G_HUMAN
ID HV3G_HUMAN STANDARD; PRT; 122 AA.
AC P01768;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE Ig heavy chain V-III region CAM.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=81013859; PubMed=6774332;
RA Lehman D.W., Putnam F.W.;
RT "Amino acid sequence of the variable region of a human mu chain;
RT location of a possible JH segment.";
RL Proc. Natl. Acad. Sci. U.S.A. 77:3239-3243(1980).
CC -1- MISCELLANEOUS: THIS MU CHAIN WAS ISOLATED FROM THE PLASMA OF A
CC PATIENT WITH MACROGLOBULINEMIA.
DR PTR; A02051; M3HUM.
```

```
DR HSSP; P01772; 2IG2.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig_1.
DR SMART; SM00406; Igv; 1.
KW Immunoglobulin V region.
FT MOD_RES 1
FT NON_TER 122 122 PYRROLIDONE CARBOXYLIC ACID.
SQ SEQUENCE 122 AA; 13668 MW; A42D0F17D252F1C2 CRC64;

Query Match 70.9%; Score 456.5; DB 1; Length 122;
Best Local Similarity 73.0%; Pred. No. 2.2e-38;
Matches 89; Conservative 13; Mismatches 19; Indels 1; Gaps 1;

QY 1 QVQLVSGGGLVQPGGSLRLSCAASGFTFRSNAMGWVRQAPGKGLVWVSIGSGGSTYY 60
DB 1 QVQLVSGGGLVQPGGSLRLSCAASGFTFRSNAMGWVRQAPGKGLVWVSIGSGGSTYY 60
QY 61 ADSVKGRFTISRDNKNTLYLQMSLRADTAIVYCAKHTGGVWDPIDYWGQGLTVTV 119
DB 61 ADSVKGRFTISRDNKNTLYLQMSLRADTAIVYCAKHTGGVWDPIDYWGQGLTVTV 120
QY 120 SS 121
DB 121 SS 122

RESULT 4
HV3A_HUMAN
ID HV3A_HUMAN STANDARD; PRT; 122 AA.
AC P01762;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE Ig heavy chain V-III region TRO.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE (MYELOMA PROTEIN TRO).
RX MEDLINE=76023781; PubMed=809331;
RA Kratzin H., Altevoigt P., Ruban E., Kortt A., Starosclik K.,
RA Hilschmann N.;
RT "The primary structure of a monoclonal IgA-immunoglobulin (IgA TRO);
RT II. The amino acid sequence of the H-chain, alpha-type, subgroup III;
RT structure of the complete IgA-molecule.";
RL Hoppe-Seyler's Z. Physiol. Chem. 356:1337-1342(1975).
CC -1- MISCELLANEOUS: THE SEQUENCE OF THE C REGION IS ALSO GIVEN.
DR PIR; A02045; ALHUTR.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig_1.
DR SMART; SM00406; Igv; 1.
KW Immunoglobulin V region.
FT MOD_RES 1
FT NON_TER 122 122 PYRROLIDONE CARBOXYLIC ACID.
SQ SEQUENCE 122 AA; 13472 MW; 2E21A11DA04D80F9 CRC64;

Query Match 70.0%; Score 450.5; DB 1; Length 122;
Best Local Similarity 68.3%; Pred. No. 8.5e-38;
Matches 84; Conservative 16; Mismatches 20; Indels 3; Gaps 2;

QY 1 QVQLVSGGGLVQPGGSLRLSCAASGFTFRSNAMGWVRQAPGKGLVWVSIGSGGSTYY 60
DB 1 QVQLVSGGGLVQPGGSLRLSCAASGFTFRSNAMGWVRQAPGKGLVWVSIGSGGSTYY 60
QY 61 ADSVKGRFTISRDNKNTLYLQMSLRADTAIVYCAKHTGGVWDPIDYWGQGLTVTV 118
DB 61 ADSVKGRFTISRDNKNTLYLQMSLRADTAIVYCAKHTGGVWDPIDYWGQGLTVTV 119
```



```
Query Match      68.1%; Score 438.5; DB 1; Length 116;
Best Local Similarity 87.8%; Pred. No. 1.2e-36;
Matches 86; Conservative 4; Mismatches 7; Indels 1; Gaps 1;

QY 1 QVQLVQSGGGLVQPQGGSLRLSCAASGFTFRSNAMGWVRQAPGKGLWVSGISGGSTYY 60
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 20 EVLVESGGGLIQPGGSLRLSCAASGFTVSSNMVSKVRQPGKGLWVSVIY-SGGSTYY 78
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 61 ADSYKGRFTISRDNKNTLYLQNSLRADETAVYYCAK 98
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 79 ADSYKGRFTISRDNKNTLYLQNSLRADETAVYYCAR 116
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 8
HV3K_HUMAN STANDARD; PRT; 122 AA.
AC P01769;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V-III region GA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=74175307; PubMed=4208843;
RA Florent G., Lehman D., Putnam F.W.;
RT "The switch point in mu heavy chains of human IgM immunoglobulins.";
RL Biochemistry 13:2482-2498(1974).
CC -!- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM A WALDENSTROM'S
CC MACROGLOBULIN.
DR PIR; A02052; M3HUGA.
DR HSSP; P01772; 2IG2.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IgV; 1.
KW Immunoglobulin V region.
FT MOD_RES 122 122 PYRROLIDONE CARBOXYLIC ACID.
FT NON_TER 122 122
SQ SEQUENCE 122 AA; 13166 MW; 74E5B6959E84100A CRC64;

Query Match      67.9%; Score 437.5; DB 1; Length 122;
Best Local Similarity 68.9%; Pred. No. 1.6e-36;
Matches 84; Conservative 16; Mismatches 21; Indels 1; Gaps 1;

QY 1 QVQLVQSGGGLVQPQGGSLRLSCAASGFTFRSNAMGWVRQAPGKGLWVSGISGGSTYY 60
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 QVQLVZSGGAVZPGRSLRLSCAASGFTFRSNAMHWVRQAPGKGLZWLVSISYGBBZY 60
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 61 ADSYKGRFTISRDNKNTLYLQNSLRADETAVYYCAKH-TGGGVMDPIDYWGQGLTVV 119
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 AASYKGRFTISRDBSKETWLENSLRADENTAVYYCARSGIALGSGVAGTDYWGZGLTVI 120
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 120 SS 121
    ||
Db 121 SS 122

RESULT 9
HV3K_HUMAN STANDARD; PRT; 126 AA.
AC P01772;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V-III region KOL.
OS Homo sapiens (Human).
```

```
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE, AND DISULFIDE BONDS.
RX MEDLINE=83289131; PubMed=6884994;
RA Schmidt W.E., Jung H.-D., Palm W., Hilschmann N.;
RT "Three-dimensional structural determination of antibodies. Primary
RT structure of crystallized monoclonal immunoglobulin IgG1 KOL, I.";
RL Hoppe-Seyler's Z. Physiol. Chem. 364:713-747(1987).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
RX MEDLINE=81072295; PubMed=7441755;
RA Marquart M., Deisenhofer J., Huber R., Palm W.;
RT "Crystallographic refinement and atomic models of the intact
RT immunoglobulin molecule KOL and its antigen-binding fragment at 3.0 A
RT and 1.0-A resolution.";
RL J. Mol. Biol. 141:369-391(1980).
DR PIR; A02055; G1HUKL.
DR PDB; 2F84; 12-JUL-89.
DR PDB; 2IG2; 12-JUL-89.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IgV; 1.
KW Immunoglobulin V region; 3D-structure.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT DISULFID 22 96
FT DISULFID 105 110
FT STRAND 3 7
FT STRAND 11 12
FT TURN 14 15
FT STRAND 18 25
FT HELIX 29 31
FT STRAND 34 39
FT TURN 41 42
FT STRAND 46 51
FT TURN 53 54
FT STRAND 58 60
FT TURN 62 67
FT STRAND 68 73
FT TURN 74 77
FT STRAND 78 83
FT HELIX 88 90
FT STRAND 92 99
FT STRAND 106 106
FT TURN 107 108
FT STRAND 109 109
FT STRAND 113 113
FT STRAND 120 124
FT NON_TER 126 126
SQ SEQUENCE 126 AA; 13718 MW; E4D71B52B16F8776 CRC64;
```

```
Query Match      67.9%; Score 437.5; DB 1; Length 126;
Best Local Similarity 69.8%; Pred. No. 1.7e-36;
Matches 88; Conservative 10; Mismatches 23; Indels 5; Gaps 1;

QY 1 QVQLVQSGGGLVQPQGGSLRLSCAASGFTFRSNAMGWVRQAPGKGLWVSGISGGSTYY 60
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 QVQLVESGGGVQPGKSLRLSCSSGFIPTSTAMTWVRQAPGKGLWVAIWDGSDQHY 60
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 61 ADSYKGRFTISRDNKNTLYLQNSLRADETAVYYCAKH-TGGGVMDPIDYWGQGLTVV 115
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 ADSYKGRFTISRDNKNTLYLQNSLRADETAVYYCAKH-TGGGVMDPIDYWGQGLTVV 120
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 116 LVTYSS 121
    :|||||
Db 121 PVTYSS 126
    :|||||

RESULT 10
HV16_MOUSE
```

```
ID HV16_MOUSE STANDARD; PRT; 136 AA.
AC P01783;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V region MOPC 21 precursor (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=81234548; PubMed=6788376;
RA Botwell A.L.M., Paskind M., Reth M., Imanishi-Kari T., Rajewsky K.,
RA Baltimore D.;
RT "Heavy chain variable region contribution to the NpB family of
RT antibodies: somatic mutation evident in a gamma 2a variable region.";
RL Cell 24:625-637(1981).
RN [2]
RP SEQUENCE OF 17-136.
RX MEDLINE=77100368; PubMed=401950;
RA Adetugbo K., Milstein C., Secher D.S.;
RT "Molecular analysis of spontaneous somatic mutants.";
RL Nature 265:299-304(1977).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: J00522; AKD15290.1; -.
DR PIR: A02086; GIM521.
DR HSP: P01772; 2F84.
DR InterPro: IPR003006; Ig_MHC.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00406; IGV; 1.
KW Immunoglobulin V region; Signal.
FT NON_TER 1
FT SIGNAL 1
FT CHAIN 16
FT DOMAIN 17 136 IG HEAVY CHAIN V REGION MOPC 21.
FT DOMAIN 115 119 D SEGMENT.
FT DOMAIN 120 136 JH4 SEGMENT.
FT DISULFID 38 112
FT CONFLICT 75 78 HYAD -> DYAH (IN REF. 2).
FT CONFLICT 89 90 DN -> ND (IN REF. 2).
FT CONFLICT 115 115 W -> H (IN REF. 2).
FT CONFLICT 120 120 Y -> W (IN REF. 2).
FT NON_TER 136 136
SQ SEQUENCE 136 AA; 2276A98DBDBF7016 CRC64;

Query Match 67.6%; Score 435.5; DB 1; Length 136;
Best Local Similarity 71.7%; Pred. No. 2.9e-36;
Matches 86; Conservative 12; Mismatches 21; Indels 1; Gaps 1;

QY 2 VOLVSGGGLVPGGSLRLSCAASGFTFSRNAMGWVROAPGKGLWVSGISGGSTYYA 61
DB 18 VOLVESGGGLVPGGSLRLSCAASGFTFSRNAMGWVROAPGKGLWVSGISGGSTYYA 77
QY 62 DSVKGRFTISRDNKNTLYLQMSLRSEDATVYCAKHTGGGVMDPTDYGQGTFLTVSS 121
DB 78 DTVKGRFTISRDNKNTLYLQMSLRSEDATVYCAKHTGGGVMDPTDYGQGTFLTVSS 136

RESULT 11
HV3D_HUMAN STANDARD; PRT; 115 AA.
ID HV3D_HUMAN
AC P01765;
DT 21-JUL-1986 (Rel. 01, Created)
```

```
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V-III region TIL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiini; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=78005528; PubMed=409716;
RA Wang A.-C., Wang I.Y., Fudenberg H.H.;
RT "Immunoglobulin structure and genetics. Identity between variable
RT regions of a mu and a gamma2 chain.";
RL J. Biol. Chem. 252:7192-7199(1977).
CC -1- MISCELLANEOUS: THE SEQUENCES OF THE V REGIONS OF THE HEAVY CHAINS
CC OF IGM AND IGG2 ISOLATED FROM A SINGLE PATIENT WITH BICLONAL
CC GAMMOPATHY ARE IDENTICAL. THEIR LIGHT CHAINS ARE APPARENTLY ALSO
CC IDENTICAL.
DR PIR: A02048; H3HUTL.
DR InterPro: IPR003006; Ig_MHC.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00406; IGV; 1.
KW Immunoglobulin V region.
FT NON_TER 115
FT SEQUENCE 115 AA; 12356 MW; 4DCC67D179F62326 MD4;

Query Match 67.5%; Score 435; DB 1; Length 115;
Best Local Similarity 71.1%; Pred. No. 2.7e-36;
Matches 86; Conservative 10; Mismatches 19; Indels 6; Gaps 2;

QY 1 OVQLVSGGGLVPGGSLRLSCAASGFTFSRNAMGWVROAPGKGLWVSGISGGSTYY 60
DB 1 EVQLVSGGGLVPGGSLRLSCAASGFTFSRNAMGWVROAPGKGLWVSGISGGSTYY 60
QY 61 ADSVKGRFTISRDNKNTLYLQMSLRSEDATVYCAKHTGGGVMDPTDYGQGTFLTVSS 120
DB 61 ABSVKGRFTISRDNKNTLYLQMSLRSEDATVYCAKHTGGGVMDPTDYGQGTFLTVSS 114
QY 121 S 121
DB 115 S 115

RESULT 12
ID HV02_CANFA STANDARD; PRT; 117 AA.
AC P01785;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V region MOO.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE OF 1-112.
RX MEDLINE=77242268; PubMed=407924;
RA Wasserman R.L., Capra J.D.;
RT "Primary structure of the variable regions of two canine
RT immunoglobulin heavy chains.";
RL Biochemistry 16:3160-3168(1977).
RN [2]
RP SEQUENCE OF 113-117.
RX MEDLINE=80077682; PubMed=117299;
RA McCumber L.J., Capra J.D.;
RT "The complete amino-acid sequence of a canine mu chain.";
RL Mol. Immunol. 16:565-570(1979).
CC -1- MISCELLANEOUS: THIS MU CHAIN WAS ISOLATED FROM MYELOMA PROTEIN.
DR PIR: A02068; MHDGMO.
DR InterPro: IPR003006; Ig_MHC.
```





OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=80020921; PubMed=114209;  
RA Steiner L.A.; Garcia Pardo A.; Margolies M.N.;  
RT "Amino acid sequence of the heavy-chain variable region of the  
crystallizable human myeloma protein Dob.";  
RL Biochemistry 18:4068-4080(1979).  
RN [21]  
RP CRYSTALLIZATION.  
RX MEDLINE=80020920; PubMed=114208;  
RA Steiner L.A.; Lopes A.D.;  
RT "The crystallizable human myeloma protein Dob has a hinge-region  
deletion.";  
RL Biochemistry 18:4054-4067(1979).  
CC -1- MISCELLANEOUS: THIS GAMMA-1 MYELOMA PROTEIN HAS A DELETION IN THE  
HINGE REGION. THERE ARE NO LIGHT-HEAVY OR INTER-HEAVY CHAIN  
DISULFIDE BONDS.  
CC FIR: A02065; GIHUDE.  
DR InterPro: IPR003006; Ig\_MHC.  
DR InterPro: IPR003596; Ig\_v.  
DR Pfam: PF00047; Ig\_1.  
DR SMART: SM00406; IgV\_1.  
KW Immunoglobulin V region.  
FT NON\_TER 120  
SQ SEQUENCE 120 AA; 13440 MW; 880DDE307C4B2627 CRC64;

Query Match 65.6%; Score 422.5; DB 1; Length 120;  
Best Local Similarity 69.1%; Pred. No. 4.8e-35;  
Matches 85; Conservative 10; Mismatches 23; Indels 5; Gaps 2;

QY 1 QVQLVQSGGGLVQPGGSLRLSCAASGFTFSNAGWVRQAPGKGLVWVSGISGSGSTYY 60  
DB 1 EVQLVSGGDLVQPGSLRLSCAASGFTFSNAGWVRQAPGKGLVWVSGISGSGSTYY 60  
QY 61 ADSVKGRFTISRDNKNTLYLQMSLRADETAVYYCAKHTGGVWDV--IDYWGQGTLT 118  
DB 61 ADSVKGRFAISRDNKNTLYLQMSLRADETAVYYCAKHTGGVWDV--IDYWGQGTLT 118  
QY 119 VSS 121  
DB 118 VSS 120

Search completed: October 4, 2002, 10:28:44  
Job time: 386 sec





```
RESULT 2
Q9UL71 ID Q9UL71 PRELIMINARY; PRT; 121 AA.
AC Q9UL71:
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE MYOSIN-REACTIVE IMMUNOGLOBULIN HEAVY CHAIN VARIABLE REGION
DE (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98277139; PubMed=9614934;
RA Wu X., Liu B., Van der Merwe P.L., Kalis N.N., Berney S.M.,
RA Young D.C.;
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
RT fetus.";
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL; AF035043; AAD56279.1; -.
DR HSSP; P01772; 2FB4.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR NON_TER 1
FT NON_TER 121
SQ SEQUENCE 121 AA; 13154 MW; 2F045CCFA5D50736 CRC64;

Query Match 78.7%; Score 507; DB 4; Length 121;
Best Local Similarity 81.0%; Pred. No. 2.8e-43;
Matches 98; Conservative 8; Mismatches 15; Indels 0; Gaps 0;

QY 1 QVQLVQSGGGLVQPGGSLRLSCAASGFTFRSNAMGWVROAPGKGLVWVSGISGSGSTYY 60
DB 1 EVQLVESGGGVVQPGGSLRLSCAASGFTFDGYAHMWVROAPGKGLVWVSLISGSGSTYY 60

QY 61 ADSVKGRFTISRDNKNTLYLQMSLRAEDTAVYYCAKHTGGVWDPIDYWGQGLTVTVSS 120
DB 61 ADSVKGRFTISRDNKNTLYLQMSLRAEDTALYYCAKGVTTIYDRFDINGQGTMTVTS 120

QY 121 S 121
DB 121 S 121

RESULT 3
Q9UL90 ID Q9UL90 PRELIMINARY; PRT; 113 AA.
AC Q9UL90:
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE MYOSIN-REACTIVE IMMUNOGLOBULIN HEAVY CHAIN VARIABLE REGION
DE (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98277139; PubMed=9614934;
RA Wu X., Liu B., Van der Merwe P.L., Kalis N.N., Berney S.M.,
RA Young D.C.;
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
RT fetus.";
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL; AF035024; AAD56260.1; -.
DR HSSP; P01772; 2FB4.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR NON_TER 1
FT NON_TER 113
SQ SEQUENCE 113 AA; 12434 MW; 0DA0348154DD6061 CRC64;

Query Match 74.5%; Score 480; DB 4; Length 116;
Best Local Similarity 80.8%; Pred. No. 1.3e-40;
Matches 97; Conservative 7; Mismatches 12; Indels 4; Gaps 2;

QY 2 QVQLVQSGGGLVQPGGSLRLSCAASGFTFRSNAMGWVROAPGKGLVWVSGISGSGSTYYA 61
DB 1 VQLVESGGGVVQPGGSLRLSCAASGFTFSYAMHWVROAPGKGLVWVAVSYDGSNKYYA 60

QY 62 DSVKGRFTISRDNKNTLYLQMSLRAEDTAVYYCAKHTGGVWDPIDYWGQGLTVTVSS 121
DB 61 DSVKGRFTISRDNKNTLYLQMSLRAEDTAMFYCA--GGGGL--GLGYWGQGLTVTVSS 116

RESULT 5
Q9UL72 ID Q9UL72 PRELIMINARY; PRT; 118 AA.
```

```
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
FT NON_TER 1
SQ SEQUENCE 113 AA; 12437 MW; ED57FDD19086D007F CRC64;

Query Match 75.0%; Score 483; DB 4; Length 113;
Best Local Similarity 78.5%; Pred. No. 6.5e-41;
Matches 95; Conservative 7; Mismatches 11; Indels 8; Gaps 1;

QY 1 QVQLVQSGGGLVQPGGSLRLSCAASGFTFRSNAMGWVROAPGKGLVWVSGISGSGSTYY 60
DB 1 EVQLVESGGGVVQPGGSLRLSCAASGFTFSYAMHWVROAPGKGLVWVAVIRYDGSNKYY 60

QY 61 ADSVKGRFTISRDNKNTLYLQMSLRAEDTAVYYCAKHTGGVWDPIDYWGQGLTVTVSS 120
DB 61 ADSVKGRFTISRDNKNTLYLQMSLRAEDTAVYYCAK-----DLNYWGQGLTVTVSS 112

QY 121 S 121
DB 113 S 113

RESULT 4
Q9UL93 ID Q9UL93 PRELIMINARY; PRT; 116 AA.
AC Q9UL93:
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE MYOSIN-REACTIVE IMMUNOGLOBULIN HEAVY CHAIN VARIABLE REGION
DE (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98277139; PubMed=9614934;
RA Wu X., Liu B., Van der Merwe P.L., Kalis N.N., Berney S.M.,
RA Young D.C.;
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
RT fetus.";
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL; AF035021; AAD56257.1; -.
DR HSSP; P01772; 2FB4.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR NON_TER 1
FT NON_TER 116
SQ SEQUENCE 116 AA; 12434 MW; 0DA0348154DD6061 CRC64;

Query Match 74.5%; Score 480; DB 4; Length 116;
Best Local Similarity 80.8%; Pred. No. 1.3e-40;
Matches 97; Conservative 7; Mismatches 12; Indels 4; Gaps 2;

QY 2 QVQLVQSGGGLVQPGGSLRLSCAASGFTFRSNAMGWVROAPGKGLVWVSGISGSGSTYYA 61
DB 1 VQLVESGGGVVQPGGSLRLSCAASGFTFSYAMHWVROAPGKGLVWVAVSYDGSNKYYA 60

QY 62 DSVKGRFTISRDNKNTLYLQMSLRAEDTAVYYCAKHTGGVWDPIDYWGQGLTVTVSS 121
DB 61 DSVKGRFTISRDNKNTLYLQMSLRAEDTAMFYCA--GGGGL--GLGYWGQGLTVTVSS 116

RESULT 5
Q9UL72 ID Q9UL72 PRELIMINARY; PRT; 118 AA.
```

```
AC Q9UL72;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE MYOSIN-REACTIVE IMMUNOGLOBULIN HEAVY CHAIN VARIABLE REGION
DE (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98277139; PubMed=9614934;
RA Wu X., Liu B., Van der Merwe P.L., Kalis N.N., Berney S.M.,
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
RT fetus.";
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL: AF035042; AAD56278.1;
DR HSSP: P01772; 2FB4.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_v.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00406; IGV; 1.
FT NON_TER 1
FT NON_TER 118
SQ SEQUENCE 118 AA; 12872 MW; B4D1A5944B2D5CCA CRC64;

Query Match 74.5%; Score 479.5; DB 4; Length 118;
Best Local Similarity 81.0%; Pred. No. 1.5e-40;
Matches 98; Conservative 6; Mismatches 14; Indels 3; Gaps 2;

QY 1 QVQLVQSGGGLVQPQGGSLRLSCAASGFTFRSNAMGWVRQAPGKGLEWVSGISGGSTYY 60
Db 1 EVQLVESGGGVVPRPGGSLRLSCAASGFTEDYGMVSWVRQAPGKGLEWVSGINWNGSGTY 59
QY 61 ADSVKGRFTISRDNKNTLYLQMSLRAEDTAVYYCAKHTGGGVWDPIDYWGQGLTVTS 120
Db 60 ADSVKGRFTISRDNKNTLYLQMSLRAEDTAVYYCARDRGEFL--FDYWGQGLTVTS 117
QY 121 S 121
Db 118 S 118

RESULT 6
QHCC1
ID Q9HCC1 PRELIMINARY; PRT; 112 AA.
AC Q9HCC1
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE SINGLE CHAIN FV (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Kikuchi M., Takeda C., Tsujimoto Y., Asada S., Nagata K.;
RT "An antibody fragment2A3 specific for native lysozyme :Isolation from a
RT human synthetic phage display library and characterization.";
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB049915; BAB16829.1;
DR HSSP: P01772; 2FB4.
DR InterPro: IPR003599; Ig.
DR InterPro: IPR003600; Ig.like.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_v.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00409; IGV; 1.
DR SMART: SM00406; IGV; 1.
GN
```

```
DR SMART: SM00410; IG_like; 1.
FT NON_TER 1
FT NON_TER 112
SQ SEQUENCE 112 AA; 12243 MW; 24F1A45EC3B84783 CRC64;

Query Match 73.2%; Score 471.5; DB 4; Length 112;
Best Local Similarity 76.9%; Pred. No. 9.1e-40;
Matches 90; Conservative 11; Mismatches 11; Indels 5; Gaps 1;

QY 1 QVQLVQSGGGLVQPQGGSLRLSCAASGFTFRSNAMGWVRQAPGKGLEWVSGISGGSTYY 60
Db 1 EVQLVESGGGVVPRPGGSLRLSCAASGFTEDYGMVSWVRQAPGKGLEWVSGINWNGSGTY 60
QY 61 ADSVKGRFTISRDNKNTLYLQMSLRAEDTAVYYCAKHTGGGVWDPIDYWGQGLTV 117
Db 61 ADSVKGRFTISRDNKNTLYLQMSLRAEDTAVYYCARRRYA----LDYWGQGLTV 112

RESULT 7
Q9UL91
ID Q9UL91 PRELIMINARY; PRT; 118 AA.
AC Q9UL91;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE MYOSIN-REACTIVE IMMUNOGLOBULIN HEAVY CHAIN VARIABLE REGION
DE (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98277139; PubMed=9614934;
RA Wu X., Liu B., Van der Merwe P.L., Kalis N.N., Berney S.M.,
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
RT fetus.";
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL: AF035023; AAD56259.1;
DR HSSP: P01772; 2FB4.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_v.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00406; IGV; 1.
FT NON_TER 1
FT NON_TER 118
SQ SEQUENCE 118 AA; 12843 MW; D0633949F2AC149D CRC64;

Query Match 73.1%; Score 470.5; DB 4; Length 118;
Best Local Similarity 77.5%; Pred. No. 1.2e-39;
Matches 93; Conservative 10; Mismatches 14; Indels 3; Gaps 1;

QY 1 QVQLVQSGGGLVQPQGGSLRLSCAASGFTFRSNAMGWVRQAPGKGLEWVSGISGGSTYY 60
Db 1 EVQLVESGGGLVQPQGGSLRLSCAASGFTFSYSNMVRQAPGKGLEWVSYISITITYY 60
QY 61 ADSVKGRFTISRDNKNTLYLQMSLRAEDTAVYYCAKHTGGGVWDPIDYWGQGLTVTS 120
Db 61 ADSVKGRFTISRDNKNTLYLQMSLRAEDTAVYYCAR--GDSEAFDIWGQGLTVTS 117

RESULT 8
Q9Y509
ID Q9Y509 PRELIMINARY; PRT; 147 AA.
AC Q9Y509;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE VH3 PROTEIN (FRAGMENT).
GN VH3.
```

OS	Homo sapiens (Human).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
OX	NCBI_TaxID=9606;	
NC	[1]	
RN	SEQUENCE FROM N. A.	
RP	MEDLINE=96071149; PubMed=7475288;	
RX	Cao J., Vescio R.A., Rettig M.B., Hong C.H., Kim A., Lee J.C.,	
RA	Lichtenstein A.K., Berenson J.R.;	
RA	"A CD10-positive subset of malignant cells is identified in multiple	
RT	myeloma using PCR with patient-specific immunoglobulin gene primers."	
RT	Leukemia 9:1948-1953(1995).	
RL	EMBL; S80860; AAD14339.1; -.	
DR	HSP; P01772; 2FB4.	
DR	InterPro; IPR003006; Ig_MHC.	
DR	InterPro; IPR003596; Ig_v.	
DR	Pfam; PF00047; Ig; 1.	
DR	SMART; SM00406; IGV; 1.	
DR	NON_TER	1
FT	SEQUENCE	147 AA; 15768 MW; 84899FCAAA7BC925C CRC64;

Query Match 71.2%; Score 458.5; DB 4; Length 147;  
Best Local Similarity 72.7%; Pred. No. 2.6e-38;  
Matches 93; Conservative 7; Mismatches 19; Indels 9; Gaps 2;

QY	1	QVQLVQSGGGLVQPGGSLRLSCAASGFTPRSNAMGWVQAQPKGLEWVWSYISGSGSTYY	60
Db	1	QVHVLSEGGGVVQPKSLRSLSCAASGFTTSTGMSWVQAQPKGLDWALLISVDGSTQYY	60
QY	61	ADSVKGRFTISRDNSKNTLYLQWNSLAEDTAVYVCAKHTGGGVNDP-----IDYWGQ	113
Db	61	AGSVKGRFTISRDNSKNTLYLQMTSLURVEDTAVYVCAK--DCNYFDSVGYGYAAGIDYWGQ	118
QY	114	GTLLTVSS	121
Db	119	GTLLTVSS	126

RESULT	9	
Q9UL84	PRELIMINARY;	PRT; 122 AA.
ID	Q9UL84	
AC	Q9UL84;	
DC	01-MAY-2000 (TReMBLrel. 13, Created)	
DT	01-MAY-2000 (TReMBLrel. 13, Last sequence update)	
DT	01-DEC-2000 (TReMBLrel. 19, Last annotation update)	
DE	MYOSIN-REACTIVE IMMUNOGLOBULIN HEAVY CHAIN VARIABLE REGION (FRAGMENT).	
DE	Os Homo sapiens (Human).	
OS	Homo sapiens (Human).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.	
OX	NCBI_TaxID=9606;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RX	MEDLINE=98277139; PubMed=9614934;	
RA	Wu X., Liu B., Van der Merwe P.L., Kallis N.N., Berney S.M.,	
RA	Young D.C.;	
RT	"Myosin-reactive autoantibodies in rheumatic carditis and normal	
RT	fetus.";	
RL	Clin. Immunol. Immunopathol. 87:184-192(1998).	

Query Match	70.38;	Score 452.5;	DB 4;	Length 122;
Best Local Similarity	74.68;	Pred. No. 8.1e-38;		

	Matches	91: Conservative	8: Mismatches	22: Indels	1: Gaps
Qy	1	QVQLVQSGGGLVQPGGSLRLISCAASGFTPSRNASMGWVRAQPK;LEWVSR;TSGSGGSTYY	60		
Db	1	EYQLVESGGGVYQFCSRLISCAASRFTPSNTGMHVRVRAQPK;LEWVAAISNDGSNKFY	60		
Qy	61	ADSVAGRTTISRDNSKNTLYLQWNSLRAEDTAVYCAKHTGG;VWLPIDYWGOGTLVTV	119		
Db	61	ADSVAGRTTIFRDNSKNMMDLQWNSLRAEDTAVYCAKDERGRAVSTFYFDYWGOGTLVTV	120		
Qy	120	SS 121			
Db	121	SS 122			
RESULT	10				
Q99KA4					
ID	Q99RA4	PRELIMINARY;	PRT: 487 AA.		

RESULT 10  
Q99KA4  
ID Q99KA

Q99RA4;  
AC 01-JUN-2001 (TRENBLrel. 17, Created)  
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)  
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update;)  
DE HYPOTHETICAL 52.6 KDA PROTEIN.  
OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus  
 OX NCB\_TaxID=10090;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RP

RA Strausberg R.;  
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases

DR	EMBL; BC004786; AAH04786.1; -
DR	HSSP; P01810; 2FEJ.
DR	InterPro; IPR003599; Ig.
DR	InterPro; IPR003597; Ig_cl.
DR	InterPro; IPR003600; Ig_like.
DR	InterPro; IPR003006; Ig_MHC.
DR	InterPro; IPR003596; Ig_v.
DR	Pfam; PF00047; Ig_4.
DR	SMART; SM00409; IG; 3.
DR	SMART; SM00407; Igc1; 3.
DR	SMART; SM00406; Igv; 1.
DR	SMART; SM00410; Ig_like; 1.
DR	PROSITE; PS00290; IG_MHC; UNKNOWN_2.
DR	Hypothetical protein.
KO	SEQUENCE 487 aa; 52554 MW; 7DC8E96DB333077B ...54;
SW	

Query Match	70.1%;	Score 451.5;	DB 11;	Length 487;
Best Local Similarity	69.4%;	Pred. No. 6.1e-37;		
Matches	86;	Conservative 14;	Mismatches 21;	Indels 3;

QY	1	QVQLVQSGGGLVQPGGSLRLISCAASGFTFRSNAMGWROAPKELFNWVSGISGSGSTYY	60
		:   :     :     :     :     :     :     :     :     :     :	
DB	20	EVQLVESGGGLVQPGGSLRLISCAASGFTFRSNAMGWROAPEKLEWATLSDGGSITY	79
		:   :     :     :     :     :     :     :     :     :     :	
QY	61	ADSVAGRTETISDRNSKNTLQLQMSRAEDTAIVYCAKHTGG--GVMDPIDYWGQGTLV	117
		:   :     :     :     :     :     :     :     :     :     :	
DB	80	PDNVAGRTETISDRNAKNLQLQMSRAEDTAIVYCAKHTGG--GVMDPIDYWGQGTLV	139
QY	118	TVSSS 121	

RESULT	11
Q96K68	
ID	Q96K6

AC	Q96K68:
DT	01-DEC-2001 (TrEMBLrel. 19, Created)
DT	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT	01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE	CDNA FLJ14473 FIS, CLONE MAMMAL001080, HIGHLY SIMILAR TO HOMO

DE SAPIENS SNC73 PROTEIN (SNC73) MRNA.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=MAMMARY GLAND;  
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,  
RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,  
RA Wagatsuma M., Hosoi T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,  
RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,  
RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,  
RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,  
RA Ninomiya K., Iwayanagi T.;  
RT "NEO human cDNA sequencing project.";  
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AK027379; BAB5072.1;  
SQ SEQUENCE 494 AA; 53088 MW; 9A1D7AEB5AE4C0E CRC64;

Query Match 69.8%; Score 449.5; DB 4; Length 494;  
Best Local Similarity 72.1%; Pred. No. 9.8e-37;  
Matches 88; Conservative 11; Mismatches 22; Indels 1; Gaps 1;

QY 1 QVOLVSGGLVQPGGSLRLSCAASGFTFRSNAMGWVRQAPKGLVWSGSGSTYY 60  
Db 20 EVQLVESGGGLVQPGGSLRLSCAASGFTFRSNAMGWVRQAPKGLVWSGSGSTYY 79  
QY 61 ADSVKGRFTISRDNKNTLYLQMSLRLEDTAIVYCAKHTGGVWDPIDYWGOGTLTV 119  
Db 80 ADSVKGRFTISRDNKNTLYLQMSLRLEDTAIVYCAKHTGGVWDPIDYWGOGTLTV 139  
QY 120 SS 121  
Db 140 SS 141

RESULT 12  
QY20E7 ID Q920E7 PRELIMINARY; PRT; 119 AA.  
AC Q920E7;  
DT 01-DEC-2001 (TREMBlrel. 19, Created)  
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE PTERIN-MIMICKING ANTI-IDIOPOPE HEAVY CHAIN VARIABLE REGION  
(FRAGMENT).  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Atkin J.D., Iape A., Jennings I.G., Horaitis O., Cotton R.G.H.;  
RT "Definition of the Idiotope of Pterin-Mimicking Antibodies Expressed  
in Mammalian Cells.";  
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF307937; AAL09421.1;  
FT NON\_TER 1  
FT NON\_TER 119  
SQ SEQUENCE 119 AA; 13025 MW; F6E904044381CA7C CRC64;

Query Match 68.9%; Score 444; DB 11; Length 119;  
Best Local Similarity 71.3%; Pred. No. 5.5e-37;  
Matches 87; Conservative 13; Mismatches 18; Indels 4; Gaps 2;

QY 1 QVOLVSGGLVQPGGSLRLSCAASGFTFRSNAMGWVRQAPKGLVWSGSGSTYY 60  
Db 1 EVQLVESGGGLVQPGGSLRLSCAASGFTFRSNAMGWVRQAPKGLVWSGSGSTYY 60  
QY 61 ADSVKGRFTISRDNKNTLYLQMSLRLEDTAIVYCAKHTGGVWDPIDYWGOGTLTV 119  
Db 80 ADSVKGRFTISRDNKNTLYLQMSLRLEDTAIVYCAKHTGGVWDPIDYWGOGTLTV 139

DE SAPIENS SNC73 PROTEIN (SNC73) MRNA.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=MAMMARY GLAND;  
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,  
RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,  
RA Wagatsuma M., Hosoi T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,  
RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,  
RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,  
RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,  
RA Ninomiya K., Iwayanagi T.;  
RT "NEO human cDNA sequencing project.";  
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AK027379; BAB5072.1;  
SQ SEQUENCE 494 AA; 53088 MW; 9A1D7AEB5AE4C0E CRC64;

Query Match 69.8%; Score 449.5; DB 4; Length 494;  
Best Local Similarity 72.1%; Pred. No. 9.8e-37;  
Matches 88; Conservative 11; Mismatches 22; Indels 1; Gaps 1;

QY 1 QVOLVSGGLVQPGGSLRLSCAASGFTFRSNAMGWVRQAPKGLVWSGSGSTYY 60  
Db 20 EVQLVESGGGLVQPGGSLRLSCAASGFTFRSNAMGWVRQAPKGLVWSGSGSTYY 79  
QY 61 ADSVKGRFTISRDNKNTLYLQMSLRLEDTAIVYCAKHTGGVWDPIDYWGOGTLTV 119  
Db 80 ADSVKGRFTISRDNKNTLYLQMSLRLEDTAIVYCAKHTGGVWDPIDYWGOGTLTV 139  
QY 120 SS 121  
Db 140 SS 141

RESULT 12  
QY20E7 ID Q920E7 PRELIMINARY; PRT; 119 AA.  
AC Q920E7;  
DT 01-DEC-2001 (TREMBlrel. 19, Created)  
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE PTERIN-MIMICKING ANTI-IDIOPOPE HEAVY CHAIN VARIABLE REGION  
(FRAGMENT).  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Atkin J.D., Iape A., Jennings I.G., Horaitis O., Cotton R.G.H.;  
RT "Definition of the Idiotope of Pterin-Mimicking Antibodies Expressed  
in Mammalian Cells.";  
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF307937; AAL09421.1;  
FT NON\_TER 1  
FT NON\_TER 119  
SQ SEQUENCE 119 AA; 13025 MW; F6E904044381CA7C CRC64;

Query Match 68.9%; Score 444; DB 11; Length 119;  
Best Local Similarity 71.3%; Pred. No. 5.5e-37;  
Matches 87; Conservative 13; Mismatches 18; Indels 4; Gaps 2;

QY 1 QVOLVSGGLVQPGGSLRLSCAASGFTFRSNAMGWVRQAPKGLVWSGSGSTYY 60  
Db 1 EVQLVESGGGLVQPGGSLRLSCAASGFTFRSNAMGWVRQAPKGLVWSGSGSTYY 60  
QY 61 ADSVKGRFTISRDNKNTLYLQMSLRLEDTAIVYCAKHTGGVWDPIDYWGOGTLTV 119  
Db 80 ADSVKGRFTISRDNKNTLYLQMSLRLEDTAIVYCAKHTGGVWDPIDYWGOGTLTV 139

Db 61 PDSVKGRFTISRDNKNTLYLQMSLRLEDTAIVYCAKHTGGVWDPIDYWGOGTLTV 117  
QY 120 SS 121  
Db 118 SA 119

RESULT 13  
QY1205 ID Q91205 PRELIMINARY; PRT; 473 AA.  
AC Q91205;  
DT 01-DEC-2001 (TREMBlrel. 19, Created)  
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE HYPOTHETICAL 51.9 KDA PROTEIN.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Strausberg R.;  
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC010327; AAH10327.1;  
KW Hypothetical protein.  
SQ SEQUENCE 473 AA; 51946 MW; CP625F008932AF12 CRC64;

Query Match 67.5%; Score 434.5; DB 11; Length 473;  
Best Local Similarity 68.9%; Pred. No. 2.9e-35;  
Matches 84; Conservative 15; Mismatches 18; Indels 5; Gaps 2;

QY 1 QVOLVSGGLVQPGGSLRLSCAASGFTFRSNAMGWVRQAPKGLVWSGSGSTYY 60  
Db 20 EVQLVESGGGLVQPGGSLRLSCAASGFTFRSNAMGWVRQAPKGLVWSGSGSTYY 79  
QY 61 ADSVKGRFTISRDNKNTLYLQMSLRLEDTAIVYCAKHTGGVWDPIDYWGOGTLTV 119  
Db 80 ADSVKGRFTISRDNKNTLYLQMSLRLEDTAIVYCAKHTGGVWDPIDYWGOGTLTV 135  
QY 120 SS 121  
Db 136 SS 137

RESULT 14  
QY1WP5 ID Q91WP5 PRELIMINARY; PRT; 479 AA.  
AC Q91WP5;  
DT 01-DEC-2001 (TREMBlrel. 19, Created)  
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE HYPOTHETICAL 51.6 KDA PROTEIN.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=COLON;  
RA Strausberg R.;  
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC013656; AAH13656.1;  
KW Hypothetical protein.  
SQ SEQUENCE 479 AA; 51603 MW; ECB2D0877748584F CRC64;

Query Match 66.1%; Score 426; DB 11; Length 479;  
Best Local Similarity 66.9%; Pred. No. 2.1e-34;  
Matches 81; Conservative 20; Mismatches 14; Indels 6; Gaps 2;

QY 1 QVOLVSGGLVQPGGSLRLSCAASGFTFRSNAMGWVRQAPKGLVWSGSGSTYY 60  
Db 80 ADSVKGRFTISRDNKNTLYLQMSLRLEDTAIVYCAKHTGGVWDPIDYWGOGTLTV 135

```

Db 20 EVQLVESGGGLVQPGGSLKVSCAASGLTFSNYAMSWVRQSPKRLKLEWAAINSNGNTYY 79
Qy 61 ADSVKGRFTISRDNKNTLYLQNSLRRAEDTAVYYCAKHTGGVWDPIYWGOGTLVTYS 120
      ::::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 80 SDTMKGRFTISRDNKNTLYLQNSLRRAEDTAVYYCVR---GGYF---DVMGAGTAVTYS 133
Qy 121 S 121
Db 134 S 134

RESULT 15
Q9UL88
ID Q9UL88 PRELIMINARY; PRT; 131 AA.
AC Q9UL88;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE MYOSIN-REACTIVE IMMUNOGLOBULIN HEAVY CHAIN VARIABLE REGION
DE (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98277139; PubMed=9614934;
RA Wu X., Liu B., Van der Merwe P.L., Kallis N.N., Berney S.M.,
RA Young D.C.;
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
RT fetus.";
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL; AF035026; AAD56262.1; -.
DR HSSP; P01810; 2FBJ.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
FT NON_TER 1
FT NON_TER 131
SQ SEQUENCE 131 AA; 14142 MW; 96E7D68E375DEA0 CRC64;

Query Match 65.7%; Score 423; DB 4; Length 131;
Best Local Similarity 64.2%; Pred. No. 7.9e-35;
Matches 88; Conservative 9; Mismatches 18; Indels 22; Gaps 3;

Qy 1 QVQLVQSGGGLVQPGGSLRLSCAASGFTFSRNAMGWVRQAPGKGLEWVSGISG--SGGST 58
      :||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 EVQLVESGGGLVQPGGSLRLSCAASGFTFSKAWMSWVRQAPGKGLEWVGRIKSKTDGGIT 60
Qy 59 YVADSVKGRFTISRDNKNTLYLQNSLRRAEDTAVYYC-----AKHTGGV 104
      || |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 61 DYAAPVKRLTISRDNKNTLYLQNSLRRAEDTAVYYCTGITMIIVITTSKRRTS--- 117
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Qy 105 WDPIDYWGOGTLVTYSS 121
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 118 ---FEYWGOGTLVTYSS 131

```

Search completed: October 4, 2002, 10:28:22  
Job time: 389 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 4, 2002, 10:26:23 ; Search time 20.62 seconds  
(without alignments)  
143.332 Million cell updates/sec

Title: US-09-822-698A-3  
Perfect score: 121  
Sequence: 1 QVLQVSGGLVQPGSLRL.....GGVWDPIYWGQGLTVSS 121

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 231628 segs, 24425594 residues

Word size : 0

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents AA:\*

- 1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	39.7	117	3	US-08-545-809A-109
2	48	39.7	131	4	US-08-983-607-28
3	47	38.8	118	2	US-08-652-816A-12
4	47	38.8	120	4	US-09-025-769B-38
5	47	38.8	120	4	US-09-025-769B-63
6	47	38.8	124	4	US-08-983-607-51
7	47	38.8	140	4	US-08-983-607-32
8	47	38.8	281	4	US-09-025-769B-178
9	46	38.0	125	2	US-08-428-197-1
10	46	38.0	125	5	PCT-US93-10555-1
11	44	36.4	95	4	US-09-043-514-2
12	44	36.4	116	3	US-08-545-809A-135
13	43	35.5	113	3	US-08-974-899-6
14	43	35.5	122	2	US-07-934-373C-21
15	43	35.5	122	3	US-08-437-642B-21
16	43	35.5	122	5	PCT-US93-07832-21
17	43	35.5	283	4	US-09-420-592A-6
18	41	33.9	124	4	US-08-983-607-47
19	40	33.1	117	4	US-08-983-607-46
20	40	33.1	117	4	US-09-025-769B-24
21	40	33.1	120	1	US-08-211-202-135
22	39	32.2	98	1	US-08-211-202-118
23	39	32.2	116	1	US-08-211-202-141
24	39	32.2	117	1	US-07-942-245-36
25	39	32.2	117	3	US-08-545-809A-115
26	39	32.2	119	1	US-08-331-398A-46
27	39	32.2	119	2	US-08-331-397B-46

28 39 32.2 119 2 US-08-759-804A-46 Sequence 46, Appl  
29 39 32.2 119 4 US-09-227-693-46 Sequence 46, Appl  
30 39 32.2 120 1 US-07-942-245-35 Sequence 35, Appl  
31 39 32.2 120 2 US-08-428-197-20 Sequence 20, Appl  
32 39 32.2 120 2 US-08-428-197-22 Sequence 22, Appl  
33 39 32.2 120 2 US-08-428-197-24 Sequence 24, Appl  
34 39 32.2 120 2 US-08-428-197-26 Sequence 26, Appl  
35 39 32.2 120 2 US-08-428-197-28 Sequence 28, Appl  
36 39 32.2 120 2 US-08-428-197-30 Sequence 30, Appl  
37 39 32.2 120 2 US-08-428-197-32 Sequence 32, Appl  
38 39 32.2 120 2 US-08-428-197-34 Sequence 34, Appl  
39 39 32.2 120 2 US-08-428-197-40 Sequence 40, Appl  
40 39 32.2 120 5 PCT-US93-10555-20 Sequence 20, Appl  
41 39 32.2 120 5 PCT-US93-10555-22 Sequence 22, Appl  
42 39 32.2 120 5 PCT-US93-10555-24 Sequence 24, Appl  
43 39 32.2 120 5 PCT-US93-10555-26 Sequence 26, Appl  
44 39 32.2 120 5 PCT-US93-10555-28 Sequence 28, Appl  
45 39 32.2 120 5 PCT-US93-10555-30 Sequence 30, Appl

## ALIGNMENTS

RESULT 1  
US-08-545-809A-109  
; Sequence 109 Application US/08545809A  
; Patent No. 6096878  
; GENERAL INFORMATION:  
; APPLICANT: Honjo, Tasuku  
; APPLICANT: Matsuda, Fumihiko  
; TITLE OF INVENTION: HUMAN IMMUNOGLOBULIN VH GENE  
; NUMBER OF SEQUENCES: 145  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Fish & Richardson, P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: US  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows95  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/545, 809A  
; FILING DATE: 27-MAR-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/JP93/00603  
; FILING DATE: 10-MAY-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Freeman, John W.  
; REGISTRATION NUMBER: 29, 066  
; REFERENCE/DOCKET NUMBER: 06501/004001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-542-5070  
; TELEFAX: 617-542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 109:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 117 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-545-809A-109

Query Match 39, 78; Score 48; DB 3; Length 117;

Best Local Similarity 100.0%; Pred. No. 6, 1e-36;

Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 ISGSGGTYADSVKGRFTISRDNSKNTLYLQMNLSRAEDTAVYICAK 98

Db 70 ISSGGSTYVADSVKGRFTISRDNKNTLYLQMSLRAEDTAVYYCAK 117  
|||||

## RESULT 2

US-08-983-607-28  
; Sequence 28, Application US/08983607  
; Patent No. 6140470  
; GENERAL INFORMATION:  
; APPLICANT: Alan Garen  
; TITLE OF INVENTION: Human Anti-Tumor Monoclonal Anti-  
; TITLE OF INVENTION: Bodies  
; NUMBER OF SEQUENCES: 51  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Department of Molecular Biophysics  
; ADDRESSEE: and Biochemistry, Yale University  
; STREET: 266 Whitney Avenue  
; CITY: New Haven  
; STATE: Connecticut  
; COUNTRY: United States of America  
; ZIP: 06520-8114  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" 1.44 Mb diskette  
; COMPUTER: IBM PC  
; OPERATING SYSTEM: MS DOS  
; SOFTWARE: Word Processing  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/983.607  
; FILING DATE: April 27, 1998  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/IB96/01032  
; FILING DATE: June 28, 1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mary M. Krinsky  
; REGISTRATION NUMBER: 32423  
; REFERENCE/DOCKET NUMBER: OCR-679  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 203-773-9544  
; TELEFAX: 203-773-1183  
; INFORMATION FOR SEQ ID NO: 28:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 131 residues  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE:  
; DESCRIPTION: polypeptide  
; ORIGINAL SOURCE:  
; ORGANISM: Homo sapiens (melanoma patient immu-  
; ORGANISM: nized with autologous tumor cells)  
; INDIVIDUAL ISOLATE: peripheral blood lymphocytes  
; IMMEDIATE SOURCE:  
; LIBRARY: DM14 scrv antibodies obtained from  
; LIBRARY: fUSE5 fusion phage construct  
; CLONE: V13  
; FEATURE:  
; NAME/KEY: heavy chain  
US-08-983-607-28

Query Match 39.7%; Score 48; DB 4; Length 131;  
Best Local Similarity 100.0%; Pred. No. 6.7e-36;  
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 51 ISSGGSTYVADSVKGRFTISRDNKNTLYLQMSLRAEDTAVYYCAK 98  
|||||

Db 51 ISSGGSTYVADSVKGRFTISRDNKNTLYLQMSLRAEDTAVYYCAK 98  
|||||

## RESULT 3

US-08-983-607-28

US-08-652-816A-12  
; Sequence 12, Application US/08652816A  
; Patent No. 5872215  
; GENERAL INFORMATION:  
; APPLICANT: Osbourn, JK  
; APPLICANT: Allen, DJ  
; TITLE OF INVENTION: Specific binding members, materials and  
; TITLE OF INVENTION: methods.  
; NUMBER OF SEQUENCES: 53  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray, Borun  
; STREET: 6300 Sears Tower, 233 South Wacker Dr  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States of America  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (BPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/652.816A  
; FILING DATE: 23-MAY-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9125579.4  
; FILING DATE: 02-DEC-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9125579.8  
; FILING DATE: 02-DEC-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9206318.9  
; FILING DATE: 24-MAR-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9206372.6  
; FILING DATE: 23-SEP-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9525004.9  
; FILING DATE: 07-DEC-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9610824.6  
; FILING DATE: 23-MAY-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/GB92/02240  
; FILING DATE: 02-DEC-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/244,597  
; FILING DATE: 01-JUN-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: David W. Clough  
; REGISTRATION NUMBER: 36,107  
; REFERENCE/DOCKET NUMBER: 28111/33308  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-474-6300  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 118 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
US-08-652-816A-12

Query Match 38.8%; Score 47; DB 2; Length 118;  
Best Local Similarity 100.0%; Pred. No. 4.8e-35;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 51 ISSGGSTYVADSVKGRFTISRDNKNTLYLQMSLRAEDTAVYYCA 97  
|||||

Db 51 ISSGGSTYVADSVKGRFTISRDNKNTLYLQMSLRAEDTAVYYCA 97  
|||||

## RESULT 4

US-09-025-769B-38

```

? COUNTRY: USA
? ZIP: 10021
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patent Release #1.0, Version #1.30 (EPO)
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/09/025,769B
? FILING DATE: 18-FEB-1998
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: EP 95 11 3021.0
? FILING DATE: 18-AUG-1995
? ATTORNEY/AGENT INFORMATION:
? NAME: James F. Haley, Jr., Esq.
? REGISTRATION NUMBER: 27,794
? REFERENCE/DOCKET NUMBER: MORPHO/5
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (212)596-9000
? TELEFAX: (212)596-9090
? INFORMATION FOR SEQ ID NO: 63:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 120 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: protein
US-09-025-769B-63

Query Match      38.8%; Score 47; DB 4; Length 120;
Best Local Similarity 100.0%; Pred. No. 4.9e-35;
Matches 47; Conservative 0; Mismatches 0; Indels

QY 51 ISGSGSTYYADSVKGRFTISRDNSKNTLYIQMNSLRAEDEVYCA 97
|||||
DB 51 ISGSGSTYYADSVKGRFTISRDNSKNTLYIQMNSLRAEDEVYCA 97
|||||

RESULT          6
US-08-983-607-51
? Sequence 51, Application US/08983607
? Patent No. 6140470
? GENERAL INFORMATION:
? APPLICANT: Alan Garen
? APPLICANT: Xiaohong Cai
? TITLE OF INVENTION: Human Anti-Tumor Monoclonal Anti-
? TITLE OF INVENTION: bodies
? NUMBER OF SEQUENCES: 51
? CORRESPONDENCE ADDRESS:
? ADDRESSED: Department of Molecular Biophysics
? ADDRESSEE: and Biochemistry, Yale University
? STREET: 266 Whitney Avenue
? CITY: New Haven
? STATE: Connecticut
? COUNTRY: United States of America
? ZIP: 06520-8114
? COMPUTER READABLE FORM:
? MEDIUM TYPE: 3.5" 1.44 Mb diskette
? COMPUTER: IBM PC
? OPERATING SYSTEM: MS DOS
? SOFTWARE: Word Processing
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/983,607
? FILING DATE: April 27, 1998
? CLASSIFICATION: 435
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: PCT/IB96/01032
? FILING DATE: June 28, 1996
? CLASSIFICATION: 435
? ATTORNEY/AGENT INFORMATION:
? NAME: Mary M. Krinsky
? REGISTRATION NUMBER: 32423
? REFERENCE/DOCKET NUMBER: OCR-679

```

```
;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 203-773-9544
; TELEFAX: 203-773-1183
; INFORMATION FOR SEQ ID NO: 51:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 124 residues
; TYPE: amino acid
; STRANDEDNESS: single
; MOLECULE TYPE:
; DESCRIPTION: polypeptide
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens (melanoma patient
; INDIVIDUAL ISOLATE: peripheral blood lympho-
; IMMEDIATE SOURCE:
; LIBRARY: VH antibodies obtained from fUSE5
; CLONE: E-13
; FEATURE:
; NAME/KEY: heavy chain
; US-08-983-607-51

Query Match 38.8%; Score 47; DB 4; Length 124;
Best Local Similarity 100.0%; Pred. No. 5e-35;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 ISGSGGTYADSVKGRFTISRDNKNTLYLQMSLRADTAAYCA 97
Db 51 ISGSGGTYADSVKGRFTISRDNKNTLYLQMSLRADTAAYCA 97

RESULT 7
US-08-983-607-32
; Sequence 32, Application US/08983607
; Patent No. 6140470
; GENERAL INFORMATION:
; APPLICANT: Alan Garen
; TITLE OF INVENTION: Human Anti-Tumor Monoclonal Anti-
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Department of Molecular Biophysics
; and Biochemistry, Yale University
; STREET: 266 Whitney Avenue
; CITY: New Haven
; STATE: Connecticut
; COUNTRY: United States of America
; ZIP: 06520-8114
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" 1.44 Mb diskette
; COMPUTER: IBM PC
; OPERATING SYSTEM: MS DOS
; SOFTWARE: Word Processing
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/983.607
; FILING DATE: April 27, 1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/IB96/01032
; FILING DATE: June 28, 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mary M. Krinsky
; REGISTRATION NUMBER: 32423
; REFERENCE/DOCKET NUMBER: OCR-679
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 203-773-9544
; TELEFAX: 203-773-1183
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 281 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
```

```
;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 140 residues
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE:
; DESCRIPTION: polypeptide
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens (melanoma patient immu-
; ORGANISM: nized with autologous tumor cells)
; INDIVIDUAL ISOLATE: peripheral blood lymphocytes
; IMMEDIATE SOURCE:
; LIBRARY: DM414 scFv antibodies obtained from
; LIBRARY: fUSE5 fusion phage construct
; CLONE: V86
; FEATURE:
; NAME/KEY: heavy chain and linker
; US-08-983-607-32

Query Match 38.8%; Score 47; DB 4; Length 140;
Best Local Similarity 100.0%; Pred. No. 5.6e-35;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 ISGSGGTYADSVKGRFTISRDNKNTLYLQMSLRADTAAYCA 97
Db 51 ISGSGGTYADSVKGRFTISRDNKNTLYLQMSLRADTAAYCA 97

RESULT 8
US-09-025-769B-178
; Sequence 178, Application US/09025769B
; Patent No. 6300064
; GENERAL INFORMATION:
; APPLICANT: Knappik, Achim
; APPLICANT: Pack, Peter
; APPLICANT: Ilag, Vic
; APPLICANT: Ge, Liming
; APPLICANT: Moroney, Simon
; APPLICANT: Plueckthun, Andreas
; TITLE OF INVENTION: Protein/(Poly)peptide libraries
; NUMBER OF SEQUENCES: 373
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: James F. Haley, Jr., Esq. c/o Fish & Neave
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10021
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.3 (FPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/025,769B
; FILING DATE: 18-FEB-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 95 11 3021.0
; FILING DATE: 18-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: James F. Haley, Jr., Esq.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: MORPHO/5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)596-9000
; TELEFAX: (212)596-9090
; INFORMATION FOR SEQ ID NO: 178:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 281 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
```

US-09-025-769B-178

Query Match 38.8%; Score 47; DB 4; Length 281;  
Best Local Similarity 100.0%; Pred. No. 9.9e-35;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 ISGSGGTYADSVKGRFTISRDNKNTLYLQMSLRADTAAYYC 97  
|||||  
Db 76 ISGSGGTYADSVKGRFTISRDNKNTLYLQMSLRADTAAYYC 122

RESULT 9

US-08-428-197-1  
; Sequence 1, Application US/08428197  
; Patent No. 5891438  
; GENERAL INFORMATION:  
; APPLICANT: SILVERMAN, GREGG J.  
; TITLE OF INVENTION: METHOD FOR STIMULATING PRODUCTION OF  
; TITLE OF INVENTION: VARIABLE REGION GENE FAMILY RESTRICTED ANTIBODIES THROUGH  
; TITLE OF INVENTION: VACCINATION WITH A B-CELL SUPERANTIGEN AND CONJUGATES  
; TITLE OF INVENTION: THEREOF  
; NUMBER OF SEQUENCES: 51  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Spensley Horn Jubas & Lubitz  
; STREET: 1880 Century Park East - Suite 500  
; CITY: Los Angeles  
; STATE: California  
; COUNTRY: USA  
; ZIP: 90067

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/428,197  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/10555  
FILING DATE: 29-OCT-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Howells, Stacy L.  
REGISTRATION NUMBER: 34,842  
REFERENCE/DOCKET NUMBER: FD-2630  
TELEPHONE: (619) 455-5100  
TELEFAX: (619) 455-5110  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 125 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
IMMEDIATE SOURCE:  
CLONE: 18/2  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: 1..125

US-08-428-197-1

Query Match 38.0%; Score 46; DB 2; Length 125;  
Best Local Similarity 100.0%; Pred. No. 4e-34;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 ISGSGGTYADSVKGRFTISRDNKNTLYLQMSLRADTAAYYC 96  
|||||  
Db 51 ISGSGGTYADSVKGRFTISRDNKNTLYLQMSLRADTAAYYC 96

RESULT 10

PCT-US93-10555-1  
; Sequence 1, Application PC/TUS9310555  
; GENERAL INFORMATION:  
; APPLICANT: SILVERMAN, GREGG J.  
; TITLE OF INVENTION: METHOD FOR STIMULATING PRODUCTION OF  
; TITLE OF INVENTION: VARIABLE REGION GENE FAMILY RESTRICTED ANTIBODIES THROUGH  
; TITLE OF INVENTION: VACCINATION WITH A B-CELL SUPERANTIGEN AND CONJUGATES  
; TITLE OF INVENTION: THEREOF  
; NUMBER OF SEQUENCES: 51  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Spensley Horn Jubas & Lubitz  
; STREET: 1880 Century Park East - Suite 500  
; CITY: Los Angeles  
; STATE: California  
; COUNTRY: USA  
; ZIP: 90067

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/10555  
FILING DATE: 29-OCT-1993  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Howells, Stacy L.  
REGISTRATION NUMBER: 34,842  
REFERENCE/DOCKET NUMBER: FD-2630  
TELEPHONE: (619) 455-5100  
TELEFAX: (619) 455-5110  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 125 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
IMMEDIATE SOURCE:  
CLONE: 18/2  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: 1..125

PCT-US93-10555-1

Query Match 38.0%; Score 46; DB 5; Length 125;  
Best Local Similarity 100.0%; Pred. No. 4e-34;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 ISGSGGTYADSVKGRFTISRDNKNTLYLQMSLRADTAAYYC 96  
|||||  
Db 51 ISGSGGTYADSVKGRFTISRDNKNTLYLQMSLRADTAAYYC 96

RESULT 11

US-09-043-514-2  
; Sequence 2, Application US/09043514A  
; Patent No. 6153745  
; GENERAL INFORMATION:  
; APPLICANT: WILLIAMS, David  
; APPLICANT: BROWN, Daniel  
; APPLICANT: ZACCOLO, Manuela C.  
; APPLICANT: GHERARDI, Ermanno  
; TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO ANTIGENESIS OF NUCLEIC  
; TITLE OF INVENTION: ACIDS  
; FILE REFERENCE: 41301/251704  
; CURRENT APPLICATION NUMBER: US/09/043,514A  
; CURRENT FILING DATE: 1998-07-06  
; EARLIER APPLICATION NUMBER: GB 9519425.4  
; EARLIER FILING DATE: 1995-09-22

Query Match 38.0%; Score 46; DB 2; Length 125;  
Best Local Similarity 100.0%; Pred. No. 4e-34;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 ISGSGGTYADSVKGRFTISRDNKNTLYLQMSLRADTAAYYC 96  
|||||  
Db 51 ISGSGGTYADSVKGRFTISRDNKNTLYLQMSLRADTAAYYC 96

EARLIER APPLICATION NUMBER: GB 9602011.0  
EARLIER FILING DATE: 1996-02-01  
EARLIER APPLICATION NUMBER: PCT/GB96/02333  
EARLIER FILING DATE: 1996-09-19  
NUMBER OF SEQ ID NOS: 5  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2  
LENGTH: 95  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Primer  
FEATURE:  
OTHER INFORMATION: Primer  
US-09-043-514-2

Query Match 36.4%; Score 44; DB 4; Length 95;  
Best Local Similarity 100.0%; Pred. No. 1.9e-32;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 54 SGGSTYYADSVKGRFTISRDNKNTLYLQMNLSRAEDTAVYYCA 97  
|||||  
DB 48 SGGSTYYADSVKGRFTISRDNKNTLYLQMNLSRAEDTAVYYCA 91

RESULT 12  
US-08-545-809A-135  
Sequence 135, Application US/08545809A  
Patent No. 6096878  
GENERAL INFORMATION:  
APPLICANT: Honjo, Tasuku  
APPLICANT: Matsuda, Fumihiko  
TITLE OF INVENTION: HUMAN IMMUNOGLOBULIN VH GENE  
TITLE OF INVENTION: SEGMENTS AND DNA FRAGMENTS CONTAINING THE SAME  
NUMBER OF SEQUENCES: 145  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson, P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: US  
ZIP: 02110-2804

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows95  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/545,809A  
FILING DATE: 27-MAR-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP93/00603  
FILING DATE: 10-MAY-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Freeman, John W.  
REGISTRATION NUMBER: 29,066  
REFERENCE/DOCKET NUMBER: 06501/004001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-542-5070  
TELEFAX: 617-542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 135:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 116 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-545-809A-135

Query Match 36.4%; Score 44; DB 3; Length 116;  
Best Local Similarity 100.0%; Pred. No. 2.3e-32;

Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 54 SGGSTYYADSVKGRFTISRDNKNTLYLQMNLSRAEDTAVYYCA 97  
|||||  
DB 72 SGGSTYYADSVKGRFTISRDNKNTLYLQMNLSRAEDTAVYYCA 115

RESULT 13  
US-08-974-899-6  
Sequence 6, Application US/08974899  
Patent No. 6037454  
GENERAL INFORMATION:  
APPLICANT: Presta, Leonard G.  
APPLICANT: Jardieu, Paula M.  
TITLE OF INVENTION: Humanized Anti-CD11a Antibodies  
NUMBER OF SEQUENCES: 24  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 1 DNA Way  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080

COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WinPatIn (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/974,899  
FILING DATE:  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/031971  
FILING DATE: 11/27/96  
ATTORNEY/AGENT INFORMATION:  
NAME: Lee, Wendy M.  
REGISTRATION NUMBER: 40,378  
REFERENCE/DOCKET NUMBER: P1014R1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650/225-1994  
TELEFAX: 650/952-9881  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 113 amino acids  
TYPE: Amino Acid  
TOPOLOGY: Linear  
US-08-974-899-6

Query Match 35.5%; Score 43; DB 3; Length 113;  
Best Local Similarity 100.0%; Pred. No. 1.7e-31;  
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 55 GGSTYYADSVKGRFTISRDNKNTLYLQMNLSRAEDTAVYYCA 97  
|||||

RESULT 14  
US-07-934-373C-21  
Sequence 21, Application US/07934373C  
Patent No. 5821337  
GENERAL INFORMATION:  
APPLICANT: Paul J. Carter  
APPLICANT: Leonard G. Presta  
TITLE OF INVENTION: Immunoglobulin Variants  
NUMBER OF SEQUENCES: 48  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 1 DNA Way  
CITY: South San Francisco  
STATE: California

; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WinPatIn (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/934,373C  
; FILING DATE: 21-Aug-1992  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US92/05126  
; FILING DATE: 15-JUN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/715272  
; FILING DATE: 14-JUN-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lee, Wendy M.  
; REGISTRATION NUMBER: 40,378  
; REFERENCE/DOCKET NUMBER: P0709P2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650/225-1994  
; TELEFAX: 650/952-9881  
; INFORMATION FOR SEQ ID NO: 21:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 122 amino acids  
; TYPE: Amino Acid  
; TOPOLOGY: Linear  
; US-07-934-373C-21

Query Match 35.5%; Score 43; DB 2; Length 122;  
Best Local Similarity 100.0%; Pred. No. 1.8e-31;  
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 55 GGSTYYADSVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYYCA 97  
DB 55 GGSTYYADSVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYYCA 97

RESULT 15  
US-08-437-642B-21  
; Sequence 21, Application US/08437642B  
; Patent No. 6054297  
; GENERAL INFORMATION:  
; APPLICANT: Paul J. Carter  
; APPLICANT: Leonard G. Presta  
; TITLE OF INVENTION: Immunoglobulin Variants  
; NUMBER OF SEQUENCES: 47  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 1 DNA Way  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WinPatIn (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/437,642B  
; FILING DATE: 09-May-1995  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/934373  
; FILING DATE: 21-AUG-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/146206  
; FILING DATE: 17-NOV-1993  
; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/US92/05126  
; FILING DATE: 15-JUN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/715272  
; FILING DATE: 14-JUN-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lee, Wendy M.  
; REGISTRATION NUMBER: 40,378  
; REFERENCE/DOCKET NUMBER: P0709P2C1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650/225-1994  
; TELEFAX: 650/952-9881  
; INFORMATION FOR SEQ ID NO: 21:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 122 amino acids  
; TYPE: Amino Acid  
; TOPOLOGY: Linear  
; US-08-437-642B-21

Query Match 35.5%; Score 43; DB 3; Length 122;  
Best Local Similarity 100.0%; Pred. No. 1.8e-31;  
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 55 GGSTYYADSVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYYCA 97  
DB 55 GGSTYYADSVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYYCA 97

Search completed: October 4, 2002, 10:30:09  
Job time: 226 sec

us-09-822-698a-3.rai

Fri Oct 4 11:00:27 2002

\_\_\_\_\_



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 4, 2002, 10:30:13 ; Search time 15.75 Seconds  
(without alignments)  
297.464 Million cell updates/sec

Title: US-09-822-698A-3

Sequence: 1 QVQLVQSGGGLVQPGLRL.....GGVWDPIYWGQGLTVTVSS 121

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 105224 seqs, 38719550 residues

Word size : 0

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	44	36.4	116	1 HV05_CARAU	P19181 carassius a
2	37	30.6	117	1 HV3C_HUMAN	P01764 homo sapien
3	23	19.0	115	1 HV3D_HUMAN	P01765 homo sapien
4	23	19.0	116	1 HV3R_HUMAN	P01779 homo sapien
5	23	19.0	119	1 HV3M_HUMAN	P01774 homo sapien
6	23	19.0	119	1 HV3N_HUMAN	P01775 homo sapien
7	23	19.0	119	1 HV3P_HUMAN	P01777 homo sapien
8	23	19.0	120	1 HV3E_HUMAN	P01766 homo sapien
9	21	17.4	117	1 HV3Q_HUMAN	P01772 homo sapien
10	20	16.5	126	1 HV3K_HUMAN	P01776 homo sapien
11	19	15.7	114	1 HV01_CANFA	P01784 canis famil
12	17	14.0	117	1 HV17_MOUSE	P01786 mus musculu
13	17	14.0	122	1 HV20_MOUSE	P01789 mus musculu
14	17	14.0	122	1 HV21_MOUSE	P01790 mus musculu
15	17	14.0	123	1 HV18_MOUSE	P01787 mus musculu
16	17	14.0	123	1 HV19_MOUSE	P01788 mus musculu
17	17	14.0	123	1 HV22_MOUSE	P01791 mus musculu
18	17	14.0	123	1 HV23_MOUSE	P01792 mus musculu
19	17	14.0	123	1 HV24_MOUSE	P01793 mus musculu
20	17	14.0	123	1 HV25_MOUSE	P01794 mus musculu
21	17	14.0	144	1 HV26_MOUSE	P01795 mus musculu
22	16	13.2	115	1 HV3F_HUMAN	P01767 homo sapien
23	16	13.2	122	1 HV3S_HUMAN	P01780 homo sapien
24	16	13.2	122	1 HV3A_HUMAN	P01762 homo sapien
25	15	12.4	121	1 HV3J_HUMAN	P01771 homo sapien
26	14	11.6	114	1 HV3B_HUMAN	P01763 homo sapien
27	14	11.6	116	1 HV1A_RABIT	P01828 oryctolagus
28	14	11.6	117	1 HV03_CARAU	P19180 carassius a
29	13	10.7	97	1 HV56_MOUSE	P18527 mus musculu
30	13	10.7	98	1 HV57_MOUSE	P18528 mus musculu
31	13	10.7	116	1 HV3Q_HUMAN	P01778 homo sapien
32	13	10.7	116	1 HV3T_HUMAN	P01781 homo sapien
33	13	10.7	117	1 HV58_MOUSE	P18529 mus musculu

34 13 10.7 117 1 HV59\_MOUSE p18530 mus musculu  
35 13 10.7 118 1 HV3V\_HUMAN p80419 homo sapien  
36 13 10.7 119 1 HV3I\_HUMAN p01770 homo sapien  
37 13 10.7 122 1 HV3G\_HUMAN p01768 homo sapien  
38 13 10.7 124 1 HV1D\_HUMAN p01760 homo sapien  
39 13 10.7 142 1 HV01\_RAT p01805 rattus norv  
40 12 9.9 116 1 HV36\_MOUSE p01806 mus musculu  
41 12 9.9 117 1 HV2B\_RABIT p01828 oryctolagus  
42 12 9.9 117 1 HV41\_MOUSE p01811 mus musculu  
43 12 9.9 117 1 HV42\_MOUSE p01812 mus musculu  
44 12 9.9 117 1 HV54\_MOUSE p18525 mus musculu  
45 12 9.9 118 1 HV39\_MOUSE p01809 mus musculu

#### ALIGNMENTS

RESULT 1  
HV05\_CARAU STANDARD; PRT; 116 AA.  
AC P19181;  
DT 01-NOV-1990 (Rel. 16, Created)  
DT 01-NOV-1990 (Rel. 16, Last sequence update)  
DE 15-JUL-1999 (Rel. 38, Last annotation update)  
DE Ig heavy chain V region 5A precursor.  
OS Carassius auratus (Goldfish).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Cypriniformes; Cyprinidae; Carassius.  
OX NCBI\_TaxID=7957;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=88144476; PubMed=3125551;  
RA Wilson M.R., Middleton D., Warr G.W.;  
RT "Immunoglobulin heavy chain variable region gene evolution: structure  
RT and family relationships of two genes and a pseudogene in a teleost  
RT fish."  
RL Proc. Natl. Acad. Sci. U.S.A. 85:1566-1570(1988).  
DR PIR; B28966; B28966.  
DR InterPro; IPR003006; Ig\_MHC.  
DR InterPro; IPR003596; Ig\_V.  
DR Pfam; PF00047; Ig\_1.  
DR SMART; SM00406; IGV\_1.  
KW Immunoglobulin V region; Signal.  
FT SIGNAL 1 19  
FT CHAIN 20 116 IG HEAVY CHAIN V REGION  
FT DOMAIN 20 49 FRAMEWORK-1.  
FT DOMAIN 50 54 COMPLEMENTARITY-DETERMINING-1.  
FT DOMAIN 55 68 FRAMEWORK-2.  
FT DOMAIN 69 84 COMPLEMENTARITY-DETERMINING-2.  
FT DOMAIN 85 116 FRAMEWORK-3.  
FT DISULFID 41 114 BY SIMILARITY.  
FT NON\_TER 116  
SQ SEQUENCE 116 AA; 12808 MW; 9C2279E2DF199B12 CRC64;

Query Match 36.4%; Score 44; DB 1; Length 116;  
Best Local Similarity 100.0%; Pred. No. 14e-37;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 54 SGGSTYYADSVKGRFTISRDNKNTLYLQMNLSRAEDTAVYYCA 97  
|||||  
DB 72 SGGSTYYADSVKGRFTISRDNKNTLYLQMNLSRAEDTAVYYCA 115

RESULT 2  
HV3C\_HUMAN STANDARD; PRT; 117 AA.  
ID HV3C\_HUMAN  
AC P01764;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE Ig heavy chain V-III region VH26 precursor.

OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN NCBI\_TaxID=9606;  
 RP [1]  
 RX MEDLINE=81101090; PubMed=6450418;  
 RA Matthysens G., Rabbitts T.H.;  
 RT "Structure and multiplicity of genes for the human immunoglobulin  
 heavy chain variable region";  
 RL Proc. Natl. Acad. Sci. U.S.A. 77:6561-6565(1980).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; J00236; AAA53516.1; -;  
 DR EMBL; M35415; AAA58735.1; -;  
 DR PIR; A02047; H3HU26.  
 DR InterPro; IPR003006; Ig\_MHC.  
 DR Pfam; PF00047; Ig; 1.  
 DR SMART; SM00406; IGV; 1.  
 KW Immunoglobulin V region; Signal.  
 FT SIGNAL 1 19  
 FT CHAIN 20 117 IG HEAVY CHAIN V-III REGION VH26.  
 FT NON\_TER 117 117  
 FT SEQUENCE 117 AA; 12582 MW; E826733F1A3CB0F1 CRC64;  
 SQ  
 Query Match 30.6%; Score 37; DB 1; Length 117;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-30;  
 Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 62 DSVKGRFTISRDNSKNTLYLQMSLRSEDYAVYYCAK 98  
 DB 81 DSVKGRFTISRDNSKNTLYLQMSLRSEDYAVYYCAK 117  
 RESULT 3  
 HV3D\_HUMAN STANDARD; PRT; 115 AA.  
 ID HV3D\_HUMAN  
 AC P01765;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Ig heavy chain V-III region IIL.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN NCBI\_TaxID=9606;  
 RP [1]  
 RX MEDLINE=78005528; PubMed=409716;  
 RA Wang A.-C., Wang I.-F., Fudenberg H.H.;  
 RT "Immunoglobulin structure and genetics. Identity between variable  
 regions of a mu and a gamma2 chain";  
 RL J. Biol. Chem. 252:7192-7199(1977).  
 CC -1- MISCELLANEOUS: THE SEQUENCES OF THE V REGIONS OF THE HEAVY CHAINS  
 OF IGM AND IGG2 ISOLATED FROM A SINGLE PATIENT WITH BICLONAL  
 GAMMOPATHY ARE IDENTICAL. THEIR LIGHT CHAINS ARE APPARENTLY ALSO  
 IDENTICAL.  
 CC  
 DR PIR; A02048; H3HUTL.  
 DR InterPro; IPR003006; Ig\_MHC.  
 DR Pfam; PF00047; Ig; 1.  
 DR SMART; SM00406; IGV; 1.  
 KW Immunoglobulin V region.  
 FT NON\_TER 115 115

SQ SEQUENCE 115 AA; 12356 MW; 4DCC67D179F62326 CRC64;  
 Query Match 19.0%; Score 23; DB 1; Length 115;  
 Best Local Similarity 100.0%; Pred. No. 2.9e-16;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 7 SGGGLVQPGGSLRLSCAASGFTF 29  
 DB 7 SGGGLVQPGGSLRLSCAASGFTF 29  
 RESULT 4  
 HV3R\_HUMAN STANDARD; PRT; 116 AA.  
 ID HV3R\_HUMAN  
 AC P01779;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Ig heavy chain V-III region TUR.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN NCBI\_TaxID=9606;  
 RP [1]  
 RX MEDLINE=74142702; PubMed=4522793;  
 RA Capra J.D., Kehoe J.M.;  
 RT "Variable region sequences of five human immunoglobulin heavy chains  
 of the VH3 subgroup: definitive identification of four heavy chain  
 hypervariable regions";  
 RL Proc. Natl. Acad. Sci. U.S.A. 71:845-848(1974).  
 CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM AN IGAL MYELOMA  
 PROTEIN.  
 CC  
 DR PIR; A02062; A1HUTU.  
 DR InterPro; IPR003006; Ig\_MHC.  
 DR Pfam; PF00047; Ig; 1.  
 DR SMART; SM00406; IGV; 1.  
 KW Immunoglobulin V region.  
 FT NON\_TER 116 116  
 FT SEQUENCE 116 AA; 12431 MW; EB705F553A963F0C CRC64;  
 Query Match 19.0%; Score 23; DB 1; Length 116;  
 Best Local Similarity 100.0%; Pred. No. 2.9e-16;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 7 SGGGLVQPGGSLRLSCAASGFTF 29  
 DB 7 SGGGLVQPGGSLRLSCAASGFTF 29  
 RESULT 5  
 HV3M\_HUMAN STANDARD; PRT; 119 AA.  
 ID HV3M\_HUMAN  
 AC P01774;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Ig heavy chain V-III region POM.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN NCBI\_TaxID=9606;  
 RP [1]  
 RX MEDLINE=75046755; PubMed=4139708;  
 RA Capra J.D., Kehoe J.M.;  
 RT "Structure of antibodies with shared idiotype: the complete sequence  
 of the heavy chain variable regions of two immunoglobulin M  
 anti-gamma globulins";  
 RL Proc. Natl. Acad. Sci. U.S.A. 71:4032-4036(1974).

```

CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM IGM WITH ANTI-GAMMA
CC GLOBULIN ACTIVITY.
DR PIR; A02057; M3HUPM.
DR HSSP; P01772; 2IG2.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; Igv; 1.
DR Immunoglobulin V region.
FT VARIANT 54
FT N -> D (PROBABLY DUE TO DEAMIDATION
FT DURING ISOLATION).
FT /FTid=VAR_003966.
FT
FT NON_TER 119
FT SEQUENCE 119 AA; 12953 MW; 2E018AF4DCB2610 CRC64;

Query Match 19.0%; Score 23; DB 1; Length 119;
Best Local Similarity 100.0%; Pred. No. 3e-16;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 SGGGLVOPGGSRLSCAASGFTF 29
Db 7 SGGGLVOPGGSRLSCAASGFTF 29

RESULT 6
HV3N_HUMAN
ID HV3N_HUMAN STANDARD; PRT; 119 AA.
AC P01775;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V-III region LAY.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RA Capra J.D., Kehoe J.M.;
RT "Structure of antibodies with shared idiotype: the complete sequence
RT of the heavy chain variable regions of two immunoglobulin M
RT anti-gamma globulins.";
RL Proc. Natl. Acad. Sci. U.S.A. 71:4032-4036(1974).
CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM IGM WITH ANTI-GAMMA
CC GLOBULIN ACTIVITY.
DR PIR; A02058; M3HULY.
DR HSSP; P01772; 2IG2.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; Igv; 1.
DR Immunoglobulin V region.
FT NON_TER 119
FT SEQUENCE 119 AA; 12858 MW; D6338098794DCF5E CRC64;

Query Match 19.0%; Score 23; DB 1; Length 119;
Best Local Similarity 100.0%; Pred. No. 3e-16;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 SGGGLVOPGGSRLSCAASGFTF 29
Db 7 SGGGLVOPGGSRLSCAASGFTF 29

RESULT 7
HV3P_HUMAN
ID HV3P_HUMAN STANDARD; PRT; 119 AA.
AC P01777;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)

```

```

DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V-III region TEI.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=74142702; PubMed=4522793;
RA Capra J.D., Kehoe J.M.;
RT "Variable region sequences of five human immunoglobulin heavy chains
RT of the VH3 subgroup: definitive identification of four heavy chain
RT hypervariable regions.";
RL Proc. Natl. Acad. Sci. U.S.A. 71:845-848(1974).
CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM AN IGG1 MYELOMA
CC PROTEIN.
DR PIR; A02060; GIHUTE.
DR HSSP; P01772; 2IG2.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; Igv; 1.
DR Immunoglobulin V region.
FT NON_TER 119
FT SEQUENCE 119 AA; 12802 MW; 7E24DC852C7290A9 CRC64;

Query Match 19.0%; Score 23; DB 1; Length 119;
Best Local Similarity 100.0%; Pred. No. 3e-16;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 SGGGLVOPGGSRLSCAASGFTF 29
Db 7 SGGGLVOPGGSRLSCAASGFTF 29

RESULT 8
HV3E_HUMAN
ID HV3E_HUMAN STANDARD; PRT; 120 AA.
AC P01766;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V-III region BRO.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=77117674; PubMed=65324;
RA Capra J.D., Hopper J.E.;
RT "Comparative studies on monotypic IgM lambda and IgG kappa from an
RT individual patient. III. The complete amino acid sequence of the VH
RT region of the IgM paraprotein.";
RL Immunohistochemistry 13:995-999(1976).
CC -1- MISCELLANEOUS: THIS CHAIN WAS OBTAINED FROM IGM ISOLATED FROM THE
CC SERUM OF A PATIENT WITH MALIGNANT LYMPHOMA OF THE WALDENSTROM
CC TYPE.
DR PIR; A02049; M3HUBW.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; Igv; 1.
DR Immunoglobulin V region.
FT NON_TER 120
FT SEQUENCE 120 AA; 13227 MW; D3F0428F7C2E6410 CRC64;

Query Match 19.0%; Score 23; DB 1; Length 120;
Best Local Similarity 100.0%; Pred. No. 3e-16;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 7 SGGGLVQPGGSLRLSCAASGFTF 29  
D5 7 SGGGLVQPGGSLRLSCAASGFTF 29

RESULT 9  
HV30\_HUMAN STANDARD; PRT; 117 AA.  
ID HV30\_HUMAN  
AC P01776;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE Ig heavy chain V-III region WAS.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
ON NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE-74142702; PubMed-4522793;  
RA Capra J.D., Kehoe J.M.;  
RT "Variable region sequences of five human immunoglobulin heavy chains  
of the Vh3 subgroup: definitive identification of four heavy chain  
hypervariable regions.";  
RL Proc. Natl. Acad. Sci. U.S.A. 71:845-848(1974).  
CC -I- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM AN IGG1 MYELOMA  
PROTEIN.  
DR PIR: A02059; G1HUKS.  
DR HSSP: P01772; 2IG2.  
DR InterPro: IPR003006; Ig\_MHC.  
DR InterPro: IPR003596; Ig\_V.  
DR Pfam: PF00047; Ig; 1.  
DR SMART: SM00406; IGV; 1.  
KW Immunoglobulin V region.  
FT NON\_TER 117  
FT SEQUENCE 117 AA; 13091 MW; 201DEF0E1E53D9BF CRC64;

Query Match 17.4%; Score 21; DB 1; Length 117;  
Best Local Similarity 100.0%; Pred. No. 3.2e-14;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 SGGGLVQPGGSLRLSCAASGFTF 27  
D5 7 SGGGLVQPGGSLRLSCAASGFTF 27

RESULT 10  
HV3K\_HUMAN STANDARD; PRT; 126 AA.  
ID HV3K\_HUMAN  
AC P01772;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE Ig heavy chain V-III region KOL.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
ON NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE, AND DISULFIDE BONDS.  
RX MEDLINE-83289131; PubMed-6884994;  
RA Schmidt W.E., Jung H.-D., Palm W., Hilschmann N.;  
RT "Three-dimensional structure determination of antibodies. Primary  
structure of crystallized monoclonal immunoglobulin IgG1 KOL, I.";  
RL Hoppe-Seyler's Z. Physiol. Chem. 364:713-747(1983).  
RN [2]  
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).  
RX MEDLINE-81072295; PubMed-7441755;  
RA Marquart M., Deisenhofer J., Huber R., Palm W.;  
RT "Crystallographic refinement and atomic models of the intact  
immunoglobulin molecule KOL and its antigen-binding fragment at 3.0 A  
and 1.0-A resolution.";

J. Mol. Biol. 141:369-391(1980).  
DR PIR: A02055; G1HUKL.  
DR PDB: 2FB4; 12-JUL-89.  
DR PDB: 2IG2; 12-JUL-89.  
DR InterPro: IPR003006; Ig\_MHC.  
DR InterPro: IPR003596; Ig\_V.  
DR Pfam: PF00047; Ig; 1.  
DR SMART: SM00406; IGV; 1.  
KW Immunoglobulin V region; 3D-structure.  
FT MOD\_RES 1  
FT DISULFID 22 96  
FT DISULFID 105 110  
FT STRAND 3 7  
FT STRAND 11 12  
FT TURN 14 15  
FT STRAND 18 25  
FT STRAND 29 31  
FT STRAND 34 39  
FT TURN 41 42  
FT STRAND 46 51  
FT TURN 53 54  
FT STRAND 58 60  
FT TURN 62 67  
FT STRAND 68 73  
FT TURN 74 77  
FT STRAND 78 83  
FT HELIX 88 90  
FT STRAND 92 99  
FT STRAND 106 106  
FT TURN 107 108  
FT STRAND 109 109  
FT STRAND 113 113  
FT STRAND 120 124  
FT NON\_TER 126  
FT SEQUENCE 126 AA; 13718 MW; E4D71B52B16F8776 CRC64;

Query Match 16.5%; Score 20; DB 1; Length 126;  
Best Local Similarity 100.0%; Pred. No. 3.5e-13;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 YADSVKGRFTISRDNKNTL 79  
D5 60 YADSVKGRFTISRDNKNTL 79

RESULT 11  
HV01\_CANFA STANDARD; PRT; 114 AA.  
ID HV01\_CANFA  
AC P01784;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE Ig heavy chain V region GOM.  
OS Canis familiaris (Dog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Carnivora; Canidae; Canis.  
ON NCBI\_TaxID=9615;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE-77242268; PubMed-407924;  
RA Wasserman R.L., Capra J.D.;  
RT "Primary structure of the variable regions of two canine  
immunoglobulin heavy chains.";  
RL Biochemistry 16:3160-3168(1977).  
CC -I- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM A MYELOMA PROTEIN.  
DR PIR: A02067; AVDSGM.  
DR InterPro: IPR003006; Ig\_MHC.  
DR InterPro: IPR003596; Ig\_V.  
DR Pfam: PF00047; Ig; 1.  
DR SMART: SM00406; IGV; 1.  
KW Immunoglobulin V region.  
FT NON\_TER 114  
FT SEQUENCE 114



OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE  
 RX MEDLINE=81054880; PubMed=6776528;  
 RA Robinson E.A., Appella E.;  
 RT "Complete amino acid sequence of a mouse immunoglobulin alpha chain  
 (MOPC 511).";  
 RL Proc. Natl. Acad. Sci. U.S.A. 77:4909-4913(1980).  
 CC -!- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM A MYELOMA PROTEIN THAT  
 BINDS PHOSPHORYLCHOLINE.  
 DR PIR: A02070; AVMS75.  
 DR HSSP; P01789; IMCP.  
 DR InterPro; IPR003006; Ig\_MHC.  
 DR InterPro; IPR003596; Ig\_V.  
 DR Pfam; PF00047; Ig; 1.  
 DR SMART; SM00406; IGV; 1.  
 KW immunoglobulin V region.  
 FT NON\_TER 122 122  
 SQ SEQUENCE 122 AA; 13652 MW; 9F4837731EA50207 CRC64;

Query Match 14.0%; Score 17; DB 1; Length 122;  
 Best Local Similarity 100.0%; Pred. No. 3.8e-10;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 7 SGGGLVQPGGSLRLSCA 23  
 |||||  
 DB 7 SGGGLVQPGGSLRLSCA 23

RESULT 15  
 HV18\_MOUSE  
 ID HV18\_MOUSE STANDARD; PRT; 123 AA.  
 AC P01787;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Ig heavy chain V regions TEPC 15/S107/HPCM1/HPCM2/HPCM3.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE (TEPC 15).  
 RX MEDLINE=76222762; PubMed=819932;  
 RA Rudikoff S., Potter M.;  
 RT "Size differences among immunoglobulin heavy chains from  
 phosphorylcholine-binding proteins";  
 RL Proc. Natl. Acad. Sci. U.S.A. 73:2109-2112(1976).  
 RN [2]  
 RP SEQUENCE FROM N.A. (H107).  
 RX MEDLINE=80199926; PubMed=6769593;  
 RA Early P., Huang H., Davis M., Calame K., Hood L.;  
 RT "An immunoglobulin heavy chain variable region gene is generated from  
 three segments of DNA: VH, D and JH.";  
 RL Cell 19:981-992(1980).  
 RN [3]  
 RP SEQUENCE (S107).  
 RX MEDLINE=76110488; PubMed=813561;  
 RA Rudikoff S., Barstad P., Potter M., Hood L.;  
 RL Unpublished results, cited by:  
 RL Hood L., Campbell J.H., Elgin S.C.R.;  
 RL Annu. Rev. Genet. 9:305-353(1975).  
 RN [4]  
 RP SEQUENCE (HPCM1; HPCM2 AND HPCM3).  
 RX MEDLINE=81197602; PubMed=7231520;  
 RA Gearhart P.J., Johnson N.D., Douglas R., Hood L.;  
 RT "IgG antibodies to phosphorylcholine exhibit more diversity than  
 their IgM counterparts";  
 RL Nature 291:29-34(1981).  
 RL Nature 291:29-34(1981).  
 CC -!- MISCELLANEOUS: ALL THOSE SEQUENCE APPEARS TO BE IDENTICAL.

CC -!- MISCELLANEOUS: THESE CHAINS WERE ISOLATED FROM MYELOMA AND  
 HYBRIDOMA PROTEINS THAT BIND PHOSPHORYLCHOLINE.  
 DR PIR: A02070; AVMS75.  
 DR HSSP; P01789; IMCP.  
 DR InterPro; IPR003006; Ig\_MHC.  
 DR InterPro; IPR003596; Ig\_V.  
 DR Pfam; PF00047; Ig; 1.  
 DR SMART; SM00406; IGV; 1.  
 KW immunoglobulin V region; Hybridoma.  
 FT NON\_TER 123 123  
 SQ SEQUENCE 123 AA; 13777 MW; 9D58086DE12F7000 CRC64;

Query Match 14.0%; Score 17; DB 1; Length 123;  
 Best Local Similarity 100.0%; Pred. No. 3.8e-10;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 7 SGGGLVQPGGSLRLSCA 23  
 |||||  
 DB 7 SGGGLVQPGGSLRLSCA 23

Search completed: October 4, 2002, 10:36:36  
 Job time: 383 sec

RECEIVED  
FBI  
OCT 4 2002

